DIVERGENT SERIES

Oxford University Press, Amen House, London E.C.4

GLASGOW NEW YORK TORONTO MELBOURNE WELLINGTON
BOMBAY CALCUTTA MADRAS KARACHI LAHORE DACCA
CAPE TOWN SALISBURY NAIROBI IBADAN ACCRA
KUALA LUMPUR HONG KONG

DIVERGENT SERIES

BY

G. H. HARDY

EMERITUS PROFESSOR OF PURE MATHEMATICS
IN THE UNIVERSITY OF CAMBRIDGE

OXFORD
AT THE CLARENDON PRESS

FIRST EDITION 1949
REPRINTED LITHOGRAPHICALLY IN GREAT BRITAIN
AT THE UNIVERSITY PRESS, OXFORD
FROM CORRECTED SHEETS OF THE FIRST EDITION
1956, 1963

Dedicated by the author to

L. S. BOSANQUET

without whose help this book

would never have been

finished

PREFACE

HARDY in his thirties held the view that the late years of a mathematician's life were spent most profitably in writing books; I remember a particular conversation about this, and though we never spoke of the matter again it remained an understanding. The level below his best at which a man is prepared to go on working at full stretch is a matter of temperament; Hardy made his decision, and while of course he continued to publish papers his last years were mostly devoted to books; whatever has been lost, mathematical literature has greatly gained. All his books gave him some degree of pleasure, but this one, his last, was his favourite. When embarking on it he told me that he believed in its value (as he well might), and also that he looked forward to the task with enthusiasm. He had actually given lectures on the subject at intervals ever since his return to Cambridge in 1931, and had at one time or another lectured on everything in the book except Chapter XIII.

The title holds curious echoes of the past, and of Hardy's past. Abel wrote in 1828: 'Divergent series are the invention of the devil, and it is shameful to base on them any demonstration whatsoever.' In the ensuing period of critical revision they were simply rejected. Then came a time when it was found that something after all could be done about them. This is now a matter of course, but in the early years of the century the subject, while in no way mystical or unrigorous, *was* regarded as sensational, and about the present title, now colourless, there hung an aroma of paradox and audacity.

J. E. LITTLEWOOD

August 1948

NOTE

PROFESSOR Hardy, who died on 1 December 1947, had sent the galleys of Chapters I–X to the press, and read the remaining galleys, before he felt unable to continue the work. Dr. H. G. Eggleston and I, who had also been reading the proofs, completed their revision in both galley and page form. Professor W. W. Rogosinski read the manuscript of Chapters I–II and XI–XII, and Miss S. M. Edmonds that of Chapter X; and I also read the book in manuscript. Dr. Eggleston checked all the references, drew up the lists of authors and definitions, and drafted the general index; and I added the note on conventions. My own task has been greatly lightened by Dr. Eggleston's help, and also by the care and consideration of the Clarendon Press.

L. S. BOSANQUET

August 1948

CONTENTS

NOTE ON CONVENTIONS

A few conventions and familiar results, not emphasized in the text, are stated here.

Stirling's Theorem

It is proved in § 13.11 that, for large real x,

$$\log \Gamma(x+1) = (x+\tfrac{1}{2})\log x - x + \tfrac{1}{2}\log 2\pi + O(x^{-1}),$$

and generally

$$\log \Gamma(x+1) = (x+\tfrac{1}{2})\log x - x + \tfrac{1}{2}\log 2\pi + \sum_{1}^{k} \frac{(-1)^{r-1}B_r}{(2r-1)\,2r} x^{-2r+1} + O(x^{-2k-1}).$$

These formulae are used freely in the earlier chapters. The second is assumed in § 6.10 for complex x (cf. Whittaker and Watson, 251–3).

Binomial Coefficients

For $n = 0, 1, \ldots$,

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!},$$

$$\binom{n+\beta}{\beta} = \frac{(\beta+1)(\beta+2)\ldots(\beta+n)}{n!} = \binom{n+\beta}{n}.$$

It follows from Stirling's theorem that, if $\beta \neq -1, -2, \ldots$, then

$$\binom{n+\beta}{\beta} = \frac{n^\beta}{\Gamma(\beta+1)} + \sum_{s=1}^{p} c_s n^{\beta-s} + O(n^{\beta-p-1}).$$

Summation Conventions

$$\sum_{\alpha}^{\beta} f(n) \quad \text{denotes} \quad \sum_{\alpha \leqslant n \leqslant \beta} f(n);$$

if $\beta < \alpha$ this is zero.

$\sum$, written without limits, usually denotes $\sum_0^\infty$, or $\sum_1^\infty$ if a term of zero rank is not defined, but other conventions are sometimes used. Conventions are given on pp. 42, 96, 131–2, 139, 162, 205, 215, 227, 239–40, 320, 350, and 372.

Differences

$$\Delta u_n = u_n - u_{n+1}, \qquad \Delta^0 u_n = u_n,$$

$$\Delta^k u_n = \Delta\, \Delta^{k-1} u_n \qquad (k = 1, 2, \ldots).$$

INTEGRATION CONVENTIONS

'Integrable in (a,b)' means 'integrable in the Lebesgue sense in (a,b)'.

All functions that occur are assumed to be measurable. Thus, if (a,b) is a finite interval, '$f = O(1)$ in (a,b)' implies 'f is integrable in (a,b)'.

$\int_0^\infty$ denotes $\lim_{X\to\infty} \int_0^X$, if this limit exists, i.e. if the integral is convergent.

$\int$, written without limits, usually denotes $\int_0^\infty$, but other conventions are sometimes used. Conventions are given on pp. 12, 50, 98, 110, 115, 135, 156, 166, 215, 235, 257, 285, 296, 327, 330, and 338.

THE CLASSES L AND L^r $(r > 0)$

'f is $L^r(a,b)$' means '[f is measurable and] $|f|^r$ is integrable in (a,b)'.

'f is L' means 'f is L^1'. Thus 'f is $L(0,\infty)$' is equivalent to '$\int_0^\infty f\,dx$ is absolutely convergent'.

CONSTANTS

Capital letters, such as $H, K, \ldots$, are used to denote numbers independent of the variables under consideration, but are not necessarily the same at each occurrence.

O, O_L, O_R, o AND $\sim$.

If $\phi > 0$, then

$$\begin{aligned} \text{'}f = O(\phi)\text{'} &\quad \text{means} \quad \text{'}|f| < H\phi\text{'}, \\ \text{'}f = O_L(\phi)\text{' [}or\ O_R(\phi)\text{]} &\quad \text{means} \quad \text{'}f > -H\phi\text{' [}or < H\phi\text{]}, \\ \text{'}f = o(\phi)\text{'} &\quad \text{means} \quad \text{'}f/\phi \to 0\text{'}, \\ \text{'}f \sim \phi\text{'} &\quad \text{means} \quad \text{'}f/\phi \to 1\text{'}.\dagger \end{aligned}$$

The symbol $\sim$ is also used for 'has the asymptotic series', 'has the Fourier series', and 'is the Fourier transform of'.

SIGN OF x

$$\operatorname{sgn} x = \begin{cases} x/|x| & (|x| \neq 0) \\ 0 & (|x| = 0). \end{cases}$$

INTEGRAL PART OF x

$[x]$ denotes the algebraically greatest integer not exceeding x.

† Here, of course, ϕ may be negative.

I

INTRODUCTION

1.1. The sum of a series. The series

$$\sum_0^\infty a_n = a_0+a_1+a_2+\ldots$$

is said to be *convergent*, to the sum s, if the 'partial sum'

$$s_n = a_0+a_1+\ldots+a_n$$

tends to a finite limit s when $n\to\infty$; and a series which is not convergent is said to be *divergent*. Thus the series

(1.1.1) $1-1+1-1+\ldots,$ (1.1.2) $1-2+3-4+\ldots,$

(1.1.3) $1-2+4-8+\ldots,$ (1.1.4) $1-1!+2!-3!+\ldots,$

(1.1.5) $1+1+1+1+\ldots,$ (1.1.6) $1+2+4+8+\ldots,$

are divergent. The series

(1.1.7) $1+e^{i\theta}+e^{2i\theta}+\ldots,$ (1.1.8) $\frac{1}{2}+\cos\theta+\cos 2\theta+\ldots,$

are divergent for all real θ, and

(1.1.9) $$\sin\theta+\sin 2\theta+\sin 3\theta+\ldots$$

is divergent except when θ is a multiple of π, when it converges to the sum 0.

The definitions of convergence and divergence are now commonplaces of elementary analysis. The ideas were familiar to mathematicians before Newton and Leibniz (indeed to Archimedes); and all the great mathematicians of the seventeenth and eighteenth centuries, however recklessly they may seem to have manipulated series, knew well enough whether the series which they used were convergent. But it was not until the time of Cauchy that the definitions were formulated generally and explicitly.

Newton and Leibniz, the first mathematicians to use infinite series systematically, had little temptation to use divergent series (though Leibniz played with them occasionally). The temptation became greater as analysis widened, and it was soon found that they were useful, and that operations performed on them uncritically often led to important results which could be verified independently. We give a few simple examples in the next section; in Ch. II we shall give others, of greater importance, from the work of the classical analysts.

1.2. Some calculations with divergent series. We know that

$$(1.2.1)\qquad 1+x+x^2+\ldots = \frac{1}{1-x}$$

if $|x| < 1$. It seems plain that, if we are to attribute a 'sum', in some sense, to the series for other x, this sum should be formally the same. For (i) it would be very inconvenient if the formula varied in different cases; (ii) we should expect the sum s to satisfy the equations

$$s = 1+x+x^2+x^3+\ldots = 1+x(1+x+x^2+\ldots) = 1+xs;$$

and (iii) the left-hand side of (1.2.1) is the result of performing the division implied by the right, so that there is certainly one sense of '$=$' with which (1.2.1) may be said to be true for all x.

(1) Let us assume then that (1.2.1) is, in some sense, true for all x (except perhaps for $x = 1$, which plainly presents special difficulties), and operate on the formula in an entirely uncritical spirit.

Putting $x = e^{i\theta}$, where $0 < \theta < 2\pi$ (so that $x \neq 1$), we obtain

$$(1.2.2)\qquad 1+e^{i\theta}+e^{2i\theta}+\ldots = (1-e^{i\theta})^{-1} = \tfrac{1}{2}+\tfrac{1}{2}i\cot\tfrac{1}{2}\theta,$$

and so

$$(1.2.3)\quad \tfrac{1}{2}+\cos\theta+\cos 2\theta+\ldots = 0, \qquad (1.2.4)\quad \sin\theta+\sin 2\theta+\ldots = \tfrac{1}{2}\cot\tfrac{1}{2}\theta,$$

for $0 < \theta < 2\pi$. Changing θ into $\theta+\pi$, we obtain

$$(1.2.5)\quad \tfrac{1}{2}-\cos\theta+\cos 2\theta-\ldots = 0, \quad (1.2.6)\quad \sin\theta-\sin 2\theta+\ldots = \tfrac{1}{2}\tan\tfrac{1}{2}\theta,$$

for $-\pi < \theta < \pi$. For $\theta = 0$, (1.2.5) gives

$$(1.2.7)\qquad 1-1+1-\ldots = \tfrac{1}{2}.$$

(2) We now differentiate (1.2.5) and (1.2.6) repeatedly with respect to θ. We thus obtain

$$(1.2.8)\qquad \sum_1^\infty (-1)^{n-1}n^{2k}\cos n\theta = 0 \quad (k = 1,2,\ldots;\ -\pi < \theta < \pi),$$

$$(1.2.9)\qquad \sum_1^\infty (-1)^{n-1}n^{2k+1}\sin n\theta = 0,$$

$$(1.2.10)\qquad \sum_1^\infty (-1)^{n-1}n^{2k}\sin n\theta = (-1)^k\left(\frac{d}{d\theta}\right)^{2k}\tfrac{1}{2}\tan\tfrac{1}{2}\theta,$$

$$(1.2.11)\qquad \sum_1^\infty (-1)^{n-1}n^{2k+1}\cos n\theta = (-1)^k\left(\frac{d}{d\theta}\right)^{2k+1}\tfrac{1}{2}\tan\tfrac{1}{2}\theta,$$

the last three formulae for $k = 0, 1,\ldots,$ $-\pi < \theta < \pi$. In particular, putting $\theta = 0$ in (1.2.8) and (1.2.11), and $\theta = \frac{1}{2}\pi$ in (1.2.9), and remembering that the Taylor's series for $\frac{1}{2}\tan\frac{1}{2}\theta$ is

$$\tfrac{1}{2}\tan\tfrac{1}{2}\theta = \sum_{k=0}^\infty \frac{2^{2k+2}-1}{(2k+2)!}B_{k+1}\,\theta^{2k+1},$$

where B_k is Bernoulli's number, we obtain

(1.2.12) $$1^{2k}-2^{2k}+3^{2k}-\ldots = 0 \quad (k = 1,2,\ldots),$$

(1.2.13) $$1^{2k+1}-2^{2k+1}+\ldots = (-1)^k\frac{2^{2k+2}-1}{2k+2}B_{k+1} \quad (k = 0,1,\ldots),$$

(1.2.14) $$1^{2k+1}-3^{2k+1}+5^{2k+1}-\ldots = 0 \quad (k = 0,1,\ldots).$$

Similarly, starting from

(1.2.15) $$e^{i\theta}-e^{3i\theta}+e^{5i\theta}-\ldots = \frac{e^{i\theta}}{1+e^{2i\theta}} = \tfrac{1}{2}\sec\theta,$$

and remembering that

$$\sec\theta = 1+\sum_{1}^{\infty}\frac{E_k\,\theta^{2k}}{(2k)!},$$

where E_k is Euler's number, we obtain (1.2.14) and also

(1.2.16) $$1^{2k}-3^{2k}+5^{2k}-\ldots = \tfrac{1}{2}(-1)^k E_k \quad (k = 1,2,\ldots).$$

We observe in passing that (1.2.13), for $k = 0$, is

(1.2.17) $$1-2+3-4+\ldots = \tfrac{1}{4},$$

which is also the result of squaring (1.2.7) by Cauchy's rule

$$(1-1+1-\ldots)(1-1+1-\ldots) = 1.1-(1.1+1.1)+(1.1+1.1+1.1)-\ldots.$$

(3) If we integrate (1.2.5) from $\theta = 0$ to $\theta = \phi$, and then write θ again for ϕ, we obtain

(1.2.18) $$\sin\theta-\tfrac{1}{2}\sin 2\theta+\tfrac{1}{3}\sin 3\theta-\ldots = \tfrac{1}{2}\theta \quad (-\pi < \theta < \pi).$$

This series is convergent. A second integration gives

(1.2.19) $$1-\cos\theta-\frac{1-\cos 2\theta}{2^2}+\frac{1-\cos 3\theta}{3^2}-\ldots = \tfrac{1}{4}\theta^2.$$

Here we may include the limits,† and $\theta = \pi$ gives

(1.2.20) $$1+\frac{1}{3^2}+\frac{1}{5^2}+\ldots = \tfrac{1}{8}\pi^2.$$

Since

$$1+\frac{1}{3^2}+\frac{1}{5^2}+\ldots = 1+\frac{1}{2^2}+\frac{1}{3^2}+\ldots-\frac{1}{2^2}-\frac{1}{4^2}-\ldots = \left(1-\frac{1}{4}\right)\left(1+\frac{1}{2^2}+\frac{1}{3^2}+\ldots\right),$$

we deduce

(1.2.21) $$1+\frac{1}{2^2}+\frac{1}{3^2}+\ldots = \tfrac{1}{6}\pi^2,$$ (1.2.22) $$1-\frac{1}{2^2}+\frac{1}{3^2}-\ldots = \tfrac{1}{12}\pi^2,$$

and so

(1.2.23) $$\cos\theta-\frac{\cos 2\theta}{2^2}+\frac{\cos 3\theta}{3^2}-\ldots = \tfrac{1}{12}\pi^2-\tfrac{1}{4}\theta^2 \quad (-\pi \leqslant \theta \leqslant \pi).$$

† The series being uniformly convergent for all θ.

Further integrations lead to the summation of $\sum(-1)^{n-1}n^{-2k}\cos n\theta$ and $\sum(-1)^{n-1}n^{-2k-1}\sin n\theta$ by means of the Bernoullian functions.

(4) Alternatively, we could, by a more daring calculation, deduce (1.2.19) from (1.2.7) and (1.2.12), arguing that

$$\sum_1^\infty (-1)^{n-1}\frac{1-\cos n\theta}{n^2} = \sum_{n=1}^\infty \frac{(-1)^{n-1}}{n^2}\sum_{k=0}^\infty (-1)^k\frac{(n\theta)^{2k+2}}{(2k+2)!}$$

$$= \sum_{k=0}^\infty \frac{(-1)^k\theta^{2k+2}}{(2k+2)!}\sum_{n=1}^\infty (-1)^{n-1}n^{2k} = \tfrac{1}{2}\theta^2(1-1+1-\ldots) = \tfrac{1}{4}\theta^2.$$

Indeed we could generalize this argument. Suppose that

$$f(\theta) = a_0+a_1\theta^2+a_2\theta^4+\ldots$$

is convergent for all θ. Then the argument suggests that

(1.2.24)

$$\sum_1^\infty (-1)^{n-1}\frac{f(n\theta)}{n^2} = \sum_{n=1}^\infty \frac{(-1)^{n-1}}{n^2}\sum_{l=0}^\infty a_l(n\theta)^{2l} = \sum_{l=0}^\infty a_l\theta^{2l}\sum_{n=1}^\infty (-1)^{n-1}n^{2l-2}$$

$$= a_0\left(\frac{1}{1^2}-\frac{1}{2^2}+\frac{1}{3^2}-\ldots\right)+a_1\theta^2(1-1+1-\ldots) = \tfrac{1}{12}a_0\pi^2+\tfrac{1}{2}a_1\theta^2.$$

This is plainly not true generally; for example, it is false when $f(\theta) = e^{-\theta^2}$: but it is true for quite extensive classes of functions. Thus, if $f(\theta)$ is the Bessel function

$$J_0(\theta) = 1-\frac{\theta^2}{2^2}+\frac{\theta^4}{2^2.4^2}-\ldots,$$

it gives

$$(1.2.25)\quad J_0(\theta)-\frac{J_0(2\theta)}{2^2}+\frac{J_0(3\theta)}{3^2}-\ldots = \tfrac{1}{12}\pi^2-\tfrac{1}{8}\theta^2 \quad (-\pi<\theta<\pi).$$

(5) From (1.2.4) we deduce

$$\sum_1^\infty \cos n\theta\sin n\phi = \tfrac{1}{4}\{\cot\tfrac{1}{2}(\phi+\theta)+\cot\tfrac{1}{2}(\phi-\theta)\} = \frac{1}{2}\frac{\sin\phi}{\cos\theta-\cos\phi},$$

and so

$$\frac{\cos m\theta-\cos m\phi}{\cos\theta-\cos\phi} = 2\sum_{n=1}^\infty \frac{\sin n\phi}{\sin\phi}\cos n\theta(\cos m\theta-\cos m\phi)$$

for any positive integral m. If we integrate this equation from $\theta = 0$

to $\theta = \pi$ (ignoring any difficulties about the range of θ over which it may be expected to be valid†), we obtain

$$(1.2.26)\qquad \int_0^{\pi} \frac{\cos m\theta - \cos m\phi}{\cos\theta - \cos\phi}\, d\theta = \pi \frac{\sin m\phi}{\sin\phi},$$

which may be verified in various ways.

(6) It follows from (1.2.4) and (1.2.6) that

$$\sin\theta + \sin 3\theta + \ldots = \tfrac{1}{2}\operatorname{cosec}\theta, \qquad \sin 2\theta + \sin 4\theta + \ldots = \tfrac{1}{2}\cot\theta.$$

If we multiply these equations by θ, integrate from $\theta = 0$ to $\theta = \frac{1}{2}\pi$, and observe that

$$\int_0^{\frac{1}{2}\pi} \theta \sin(2n+1)\theta\, d\theta = \frac{(-1)^n}{(2n+1)^2}, \qquad \int_0^{\frac{1}{2}\pi} \theta \sin 2n\theta\, d\theta = (-1)^{n-1}\frac{\pi}{4n},$$

we obtain

$$(1.2.27)\qquad \int_0^{\frac{1}{2}\pi} \frac{\theta}{\sin\theta}\, d\theta = 2\left(1 - \frac{1}{3^2} + \frac{1}{5^2} - \ldots\right),$$

$$(1.2.28)\qquad \int_0^{\frac{1}{2}\pi} \theta\cot\theta\, d\theta = \tfrac{1}{2}\pi\log 2.$$

These formulae also may be verified independently.

1.3. First definitions. The results of the formal calculations of § 1.2 are correct wherever they can be checked: thus all of the formulae (1.2.18)–(1.2.23), (1.2.25), and (1.2.26)–(1.2.28) are correct. It is natural to suppose that the other formulae will prove to be correct, and our transformations justifiable, if they are interpreted appropriately. We should then be able to regard the transformations as shorthand representations of more complex processes justifiable by the ordinary canons of analysis. It is plain that the first step towards such an interpretation must be some definition, or definitions, of the 'sum' of an infinite series, more widely applicable than the classical definition of Cauchy.

This remark is trivial now: it does not occur to a modern mathematician that a collection of mathematical symbols should have a 'meaning' until one has been assigned to it by definition. It was not a triviality even to the greatest mathematicians of the eighteenth century. They had not the habit of definition: it was not natural to

† We have to expect trouble with (1.2.4) for $\theta = 0$ or $\theta = 2\pi$, since the left-hand side vanishes identically and the right-hand side has an infinity, and here for values of θ for which $\cos\theta = \cos\phi$. But it is not unreasonable to suppose that these difficulties will disappear when we multiply by the factor $\cos m\theta - \cos m\phi$; and the result seems to justify our expectation.

them to say, in so many words, 'by X we *mean* Y'. There are reservations to be made, to which we shall return in §§ 1.6–7; but it is broadly true to say that mathematicians before Cauchy asked not 'How shall we *define* $1-1+1-...$?' but 'What *is* $1-1+1-...$?', and that this habit of mind led them into unnecessary perplexities and controversies which were often really verbal.

It is easy now to pick out one cause which aggravated this tendency, and made it harder for the older analysts to take the modern, more 'conventional', view. It generally seems that there is only one sum which it is 'reasonable' to assign to a divergent series: thus all 'natural' calculations with the series (1.1.1) seem to point to the conclusion that its sum should be taken to be $\frac{1}{2}$. We can devise arguments leading to a different value,† but it always seems as if, when we use them, we are somehow 'not playing the game'.

The reason for this is fairly obvious. The simplest argument for (1.2.7) is '$s = 1-1+1-... = 1-(1-1+1-...) = 1-s$, and so $s = \frac{1}{2}$': we thus obtain the value $\frac{1}{2}$, whatever our definition, provided only that it satisfies certain very natural conditions.

Let us suppose, for example, that we have given any definition of the sum of a series which satisfies the following axioms:

(*A*) *if* $\sum a_n = s$ *then* $\sum ka_n = ks$;

(*B*) *if* $\sum a_n = s$ *and* $\sum b_n = t$, *then* $\sum (a_n+b_n) = s+t$;

(*C*) *if* $a_0+a_1+a_2+... = s$ *then* $a_1+a_2+a_3+... = s-a_0$, *and conversely*.

Actually, all definitions which we shall use satisfy (*A*) and (*B*), and most, though not all, satisfy (*C*). Then, if $1-1+1-... = s$, we have

$$s = 1-1+... = 1+(-1+1-...) = 1-(1-1+...) = 1-s:$$

here we have used only (*A*) and (*C*). Similarly, if $1-2+3-4+... = s$, we have

$$\begin{aligned} s = 1-2+3-... &= 1+(-2+3-4+...) = 1-(2-3+4-...) \\ &= 1-(1-1+1-...)-(1-2+3-...) = 1-\tfrac{1}{2}-s, \end{aligned}$$

and so $s = \frac{1}{4}$, in agreement with (1.2.17). Here we have used all of (*A*), (*B*), and (*C*).

We pick out here four of the large number of useful definitions which we shall have occasion to use later. We shall make systematic use of the following notations. If we define the sum of $\sum a_n$, in some new

† See § 1.6(2).

sense, say the 'Pickwickian' sense, as s, we shall say that $\sum a_n$ is *summable* (P), call s the P *sum* of $\sum a_n$, and write

$$\sum a_n = s \ \ (\mathrm{P}).$$

We shall also say that s is the P *limit* of the partial sum s_n, and write

$$s_n \to s \ \ (\mathrm{P}).$$

Our choice of letters to be associated with different definitions will be determined mainly by convenience, but sometimes also by historical considerations.

(1) If $s_n = a_0+a_1+\ldots+a_n$ and

(1.3.1) $$\lim_{n\to\infty} \frac{s_0+s_1+\ldots+s_n}{n+1} = s,$$

then we call s the (C, 1) sum of $\sum a_n$ and the (C, 1) limit of s_n.

(2) If $\sum a_n x^n$ is convergent for $0 \leqslant x < 1$ (and so for all x, real or complex, with $|x| < 1$), $f(x)$ is its sum, and

(1.3.2) $$\lim_{x\to 1-0} f(x) = s,$$

then we call s the A sum of $\sum a_n$.

(3) If $\sum a_n x^n$ is convergent for small x, and defines a function $f(x)$ of the complex variable x, one-valued and regular in an open and connected region containing the origin and the point $x = 1$; and $f(1) = s$; then we call s the $\mathfrak{E}$ sum of $\sum a_n$. The value of s may naturally depend on the region chosen.

(4) Our fourth definition requires a little more explanation. Suppose that the series $\sum a_n x^n$ converges for small x, and that

(1.3.3) $$x = \frac{y}{1-y}, \qquad y = \frac{x}{1+x},$$

so that $y = \frac{1}{2}$ corresponds to $x = 1$. Then, for small x and y, we have

$$xf(x) = \sum_0^\infty a_n x^{n+1} = a_0\frac{y}{1-y}+a_1\frac{y^2}{(1-y)^2}+a_2\frac{y^3}{(1-y)^3}+\ldots$$
$$= \sum_{p=0}^\infty a_p \sum_{m=0}^\infty \binom{p+m}{m} y^{p+m+1} = \sum_{p=0}^\infty a_p \sum_{n=p}^\infty \binom{n}{n-p} y^{n+1}.$$

Inverting the order of summation, we find that

$$xf(x) = \sum_{n=0}^\infty y^{n+1} \sum_{p=0}^n \binom{n}{n-p} a_p = \sum_{n=0}^\infty y^{n+1} \sum_{p=0}^n \binom{n}{p} a_p = \sum_{n=0}^\infty b_n y^{n+1},$$

for small y, where

(1.3.4) $$b_0 = a_0, \ \ b_n = a_0+\binom{n}{1}a_1+\binom{n}{2}a_2+\ldots+a_n.$$

If the y-series is convergent for $y = \frac{1}{2}$, to sum s, i.e. if

$$\frac{1}{2}b_0 + \frac{1}{4}b_1 + \frac{1}{8}b_2 + \ldots = \sum 2^{-n-1}b_n = s, \tag{1.3.5}$$

then we call s the (E, 1) sum of $\sum a_n$.

The letters $\mathfrak{E}$ and E both stand for Euler, A for Abel, and C for Cesàro. The reasons for these choices, and for the figures in (C, 1) and (E, 1), will appear later.

The '(C, 1)' definition was used by D. Bernoulli in 1771, but only in the special case when the series is a periodic oscillating series, i.e. when $a_{n+p} = a_n$ for a fixed p, and

$$a_0 + a_1 + \ldots + a_{p-1} = 0.$$

It had been applied to the special series (1.1.1) by Leibniz as early as 1713. But neither Leibniz nor Bernoulli said in so many words that they were giving a definition. In modern times it was used implicitly by Frobenius and Hölder in 1880 and 1882; but it does not seem to have been stated formally as a definition until 1890, when Cesàro published a paper on the multiplication of series in which, for the first time, a 'theory of divergent series' is formulated explicitly. 'Lorsque s_n, sans tendre vers une limite, admet une valeur moyenne s finie et déterminée [i.e. when (1.3.1) is true] nous dirons que la série $a_0 + a_1 + a_2 + \ldots$ est simplement indéterminée, et *nous conviendrons de dire que s est la somme de la série.*' Cesàro goes on to consider series 'r-fois indéterminées', and proves a general theorem† which will be prominent in Ch. X. Cesàro's paper has become famous, and his language now seems almost absurdly modest: 'il résulte de là une classification des séries indéterminées, qui est sans doute incomplète et pas assez naturelle . . .' In fact his classification is entirely natural.

The 'A' definition is sometimes called the 'P' definition, after Poisson, who used it, in effect, for the summation of Fourier series. It also can be traced through Euler back to Leibniz. The justification for the 'A', which is usual with English writers, lies in Abel's theorem on the continuity of power series, which establishes the 'regularity' (§ 1.4) of the method, and will be proved, as a special case of a much more general theorem, in Ch. IV.

The $\mathfrak{E}$ method embodies, in modern language, Euler's famous principle 'summa cujusque seriei est valor expressionis illius finitae, ex cujus evolutione illa series oritur'. We shall have more to say about this in §§ 1.6–7: for the moment we observe only that Euler was obviously thinking in terms of power series, and that no mathematician of his period could possibly have expressed himself on such a subject without very serious ambiguity.

Finally the '(E, 1)' method is derived from 'Euler's transformation', which was primarily a weapon for transforming slowly convergent into rapidly convergent series, but which he applied to divergent series also.

It is plain that all these methods satisfy our axiomatic requirements (A) and (B), and it is easy to verify that the first three also satisfy (C), provided that the $\mathfrak{E}$ method is associated with a definite region of continuation. We denote the partial sums of $a_0 + a_1 + a_2 + \ldots$ and $a_1 + a_2 + a_3 + \ldots$ by s_n and t_n, so that $t_n = s_{n+1} - a_0$; and write

$$f_1(x) = a_1 + a_2 x + a_3 x^2 + \ldots,$$

so that $xf_1(x) = f(x) - a_0$.

† Theorem 41.

(1) If $a_0+a_1+a_2+...$ is summable (C, 1) to s, then

$$\frac{t_0+t_1+...+t_n}{n+1} = \frac{n+2}{n+1}\left(\frac{s_0+s_1+...+s_{n+1}}{n+2}-a_0\right) \to s-a_0,$$

so that $a_1+a_2+...$ is summable (C, 1) to $s-a_0$,

(2) If $a_0+a_1+...$ is summable (A) to s, then

$$f_1(x) = x^{-1}\{f(x)-a_0\} \to s-a_0,$$

and $a_1+a_2+...$ is summable (A) to $s-a_0$.

(3) If $f(x)$ is one-valued and regular in a region including 0 and 1, and $f(1) = s$, then $f_1(x)$ is also one-valued and regular in the region, and $f_1(1) = s-a_0$.

Thus the direct statement in (C) is true of each of the three methods, and the arguments are plainly reversible. It is less obvious that the (E, 1) method satisfies (C), and we postpone the proof to § 8.3. If we take this for granted for the moment, then it becomes plain that all four methods, if they sum (1.1.1), must give the sum $\frac{1}{2}$. It is easy to verify this directly, since s_n is 1 for even and 0 for odd n, so that $s_0+s_1+...+s_n$ is $\frac{1}{2}(n+2)$ or $\frac{1}{2}(n+1)$; since $f(x) = (1+x)^{-1}$; and since $b_0 = 1$ and $b_n = 0$ for $n > 0$.

We shall see later (or the reader may verify as an exercise) that all four methods also yield the equations (1.2.2)–(1.2.7), and that the last three yield all of (1.2.8)–(1.2.17). The (C, 1) method fails with (1.2.17), since the values of s_n are $1, -1, 2, -2, 3,...$ and $s_0+s_1+...+s_n$ is $\frac{1}{2}(n+2)$ for even and 0 for odd n. It will be observed that in this case a repetition of the averaging process would give the limit $\frac{1}{4}$.

Methods (1), (2), and (4) give ∞ as the sum of (1.1.5): in the last case $b_0 = 1$, $b_1 = 2$, $b_2 = 4,...,$ so that the series (1.3.5) is $\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+....$ Method (3) is inapplicable, since $f(x) = (1-x)^{-1}$ is not regular at $x = 1$.

Methods (1) and (2) fail for (1.1.3): the values of s_n are $1, -1, 3, -5, 11,...$; and $\sum a_n x^n$ is not convergent when $x \geqslant \frac{1}{2}$. Method (3) gives the sum $\frac{1}{3}$. In method (4), $b_n = (1-2)^n = (-1)^n$, and so

$$\tfrac{1}{2}b_0+\tfrac{1}{4}b_1+\tfrac{1}{8}b_2+... = \tfrac{1}{2}-\tfrac{1}{4}+\tfrac{1}{8}-... = \tfrac{1}{3},$$

so that this method also gives $\frac{1}{3}$. This is plainly the 'right' sum, since it satisfies $s = 1-2s$.

It is also instructive to consider (1.1.6). Here method (1) gives ∞. Method (2) is inapplicable for the same reason as in the last paragraph. Method (3) gives $(1-2.1)^{-1} = -1$. Finally, with method (4), we have $b_n = (1+2)^n = 3^n$,

$$\tfrac{1}{2}b_0+\tfrac{1}{4}b_1+\tfrac{1}{8}b_2+... = \tfrac{1}{2}+\tfrac{3}{4}+...+\tfrac{1}{2}(\tfrac{3}{2})^n+...,$$

which diverges to ∞, so that the method gives ∞. It will be observed that in this case there are two suggested 'sums', viz. ∞ and -1, and that the second has an air of paradox, since it does not seem natural to attribute a negative sum to a series of positive terms.

1.4. Regularity of a method. It is easy to state in general terms some of the qualities required for a useful method of summation of divergent series. It should be *simple*, as, for example, the first two methods of § 1.3 are simple; and it should be reasonably *general*, in the sense of being applicable to a good variety of important series. There is another requirement which can be stated more exactly, that of *consistency* or *regularity*.

A method will be said to be *regular* if it sums every convergent series to its ordinary sum. Thus the (C, 1) and A methods are regular, since $\sum a_n = s$ implies both

$$S_n = \frac{s_0+s_1+...+s_n}{n+1} \to s$$

and $f(x) = \sum a_n x^n \to s$, the first by a well-known theorem of Cauchy, the second by Abel's theorem on power series.

These methods are regular in an extended sense. If a_n is real and $s_n \to \infty$ (for example, if $\sum a_n$ is a divergent series of positive terms), then $S_n \to \infty$, and the (C, 1) method gives $s = \infty$. For the A method there are two possibilities. Either $\sum a_n x^n$ diverges for some $x = x_0 < 1$, in which case it necessarily diverges to ∞ in the interval $(x_0, 1)$, and $f(x) = \infty$ in such an interval; or $\sum a_n x^n$ converges for $0 \leqslant x < 1$, in which case $f(x) \to \infty$ when $x \to 1$. In either case we can say that the A method gives $s = \infty$. When a regular method has this additional property, we shall say that it is *totally regular*. We shall see (§§ 3.6 and 4.6) that the (E, 1) method is also totally regular. It is obvious that the $\mathfrak{E}$ method is not totally regular, since it sums $1+2+4+8+...$ to -1. In fact it is not even regular, since $f(x)$ need not be regular at $x = 1$ when $\sum a_n$ converges.

1.5. Divergent integrals and generalized limits of functions of a continuous variable. It is natural to give similar definitions applying to functions of a continuous variable x. Suppose that $a(t)$ is integrable in every finite interval $(0, x)$, that

$$s(x) = \int_0^x a(t)\,dt,$$

and that we have given some 'Pickwickian' definition of the limit s of $s(x)$ when $x \to \infty$, or, what is the same thing, of

$$\int_0^\infty a(x)\,dx.$$

Then we shall say that $a(x)$ is P *integrable* in $(0,\infty)$, that $s(x)$ has the P *limit* s, that $s(x) \to s$ (P), and that

$$\int_0^\infty a(x)\,dx = s \quad (\mathrm{P}).$$

Thus the definitions corresponding to the first three of § 1.3 are as follows.

(1) If

$$\frac{1}{x}\int_0^x s(t)\,dt \to s \tag{1.5.1}$$

or, what is the same thing, if

$$\frac{1}{x}\int_0^x (x-t)a(t)\,dt \to s,$$

we shall say that

$$\int_0^\infty a(x)\,dx = s \quad (\mathrm{C},1). \tag{1.5.2}$$

(2) If

$$f(w) = \int_0^\infty e^{-wx}a(x)\,dx \tag{1.5.3}$$

is convergent for $w > 0$, and $f(w) \to s$ when $w \to +0$, we shall say that

$$\int_0^\infty a(x)\,dx = s \quad (\mathrm{A}). \tag{1.5.4}$$

(3) If there is a function $f(w)$ of the complex variable w, defined by (1.5.3) for large positive w, and one-valued and regular in an open and connected region containing the origin and the distant part of the positive real axis; and if $f(0) = s$; then we shall say that

$$\int_0^\infty a(x)\,dx = s \quad (\mathfrak{E}).\dagger \tag{1.5.5}$$

We can modify all these definitions, if we please, by a change in the lower limit. There is no useful analogue of the (E, 1) definition.

† The definition does not correspond exactly to the '$\mathfrak{E}$' definition of § 1.3, since $f(w)$ will not usually be regular at infinity.

Thus if $a(x) = e^{mix}$, where $m > 0$, we have

$$s(x) = \frac{i}{m}(1-e^{mix}), \qquad \frac{1}{x}\int_0^x s(t)\,dt = \frac{i}{m}+\frac{1-e^{mix}}{m^2x} \to \frac{i}{m},$$

so that

(1.5.6) $$\int_0^\infty e^{mix}\,dx = \frac{i}{m}, \qquad \int_0^\infty \cos mx\,dx = 0, \qquad \int_0^\infty \sin mx\,dx = \frac{1}{m}$$

all (C, 1); and $$\int_0^\infty e^{-(w-mi)x}\,dx = \frac{1}{w-mi} \to \frac{i}{m},$$

so that the A and $\mathfrak{E}$ methods give the same results. Also

$$s(x) = im^{-1}(1-e^{mix}) \to im^{-1} \quad (\mathrm{P}),$$

so that

(1.5.7) $$e^{mix} \to 0, \qquad \cos mx \to 0, \qquad \sin mx \to 0 \quad (\mathrm{P}),$$

where P may be (C, 1), A, or $\mathfrak{E}$. We have thus defined various senses in which '$\cos\infty = 0$ and $\sin\infty = 0$'.

It will be observed that

$$\frac{1}{x}\int_0^x \begin{matrix}\cos^2\\ \sin^2\end{matrix} mt\,dt = \frac{1}{2}\pm\frac{\sin 2mx}{4mx} \to \frac{1}{2},$$

so that $$\cos^2 mx \to \tfrac{1}{2}, \qquad \sin^2 mx \to \tfrac{1}{2} \quad (C, 1);$$

and it is easy to show that the A and $\mathfrak{E}$ methods give the same limits. It is not to be expected that the P limit of the square of a function should usually be the square of its P limit.

We add some examples of formal calculations with integrals analogous to those of § 1.2. All the integrations are over $(0,\infty)$. Differentiation of (1.5.6) with respect to m gives

(1.5.8) $$\int x^{2p}\cos mx\,dx = 0, \qquad \int x^{2p+1}\sin mx\,dx = 0,$$

(1.5.9) $$\int x^{2p}\sin mx\,dx = (-1)^p\frac{(2p)!}{m^{2p+1}},$$

$$\int x^{2p+1}\cos mx\,dx = (-1)^{p+1}\frac{(2p+1)!}{m^{2p+2}}.$$

If $$\phi(x) = a_0+a_1x^2+a_2x^4+\ldots,$$

and we integrate term by term, then we obtain

$$(1.5.10)\quad \int \phi(x)\cos mx\,dx = a_0 \int \cos mx\,dx + a_1 \int x^2 \cos mx\,dx + \ldots = 0,$$

$$(1.5.11)\quad \int \phi(x)\sin mx\,dx = a_0 \int \sin mx\,dx + \ldots = \frac{a_0}{m} - \frac{2!\,a_1}{m^3} + \frac{4!\,a_2}{m^5} - \ldots.$$

As is to be expected, these formulae are sometimes correct and sometimes not. Thus if $\phi(x) = J_0(x)$ they are

$$\int J_0(x)\cos mx\,dx = 0, \qquad \int J_0(x)\sin mx\,dx = \frac{1}{m} + \frac{1}{2}\frac{1}{m^3} + \ldots = \frac{1}{\sqrt{(m^2-1)}},$$

and they are correct when $m > 1$. But they are false when $m < 1$, and (1.5.10) is obviously false when $\phi(x) = e^{-x^2}$.

1.6. Some historical remarks. In the next chapter we shall give substantial examples of the use of divergent series by Euler and other early analysts. It will be convenient to lead up to them by a few more miscellaneous remarks.

(1) The earliest analysts were, on the whole, rather severely 'orthodox': their work had the arithmetical spirit of that of the Greeks. What is lacking in the work of Cavalieri, Wallis, Brouncker, Gregory (who first used the word 'convergent'), and Mercator is not *rigour* but *technique*. In particular they were handicapped by the lack of serviceable criteria for convergence. Newton was the first analyst who was the master of a really powerful technique: he regarded infinite series primarily as a tool for quadratures, and there was so much for him to do in this field that the rewards of orthodoxy were sufficient. He was no doubt aware that many of his formulae could be interpreted in different senses, for example, that

$$(1.6.1)\qquad \frac{f(x)}{g(x)} = a_0 + a_1 x + a_2 x^2 + \ldots,$$

where f and g are polynomials, could be interpreted either in the arithmetical sense which demands convergence or in the algebraical sense in which it means that

$$(1.6.2)\qquad f(x) - (a_0 + a_1 x + \ldots + a_n x^n) g(x)$$

is divisible by x^{n+1} for every value of n.

(2) Thus there is little about divergent series before Euler except in certain passages in the correspondence of Leibniz and the Bernoullis; and the impression which these leave is that Leibniz missed a great opportunity. He was on the track of at least one of the standard definitions, but gave way to the temptation of seasoning the discussion with metaphysics. The sum of $1-1+1-\ldots$ is to be $\frac{1}{2}$ on grounds of

'probability': 'porro hoc argumentandi genus, etsi Metaphysicum magis quam Mathematicum videatur, tamen firmum est: et alioqui Canonum Verae Metaphysicae major est usus in Mathesi, in Analysi, in ipsa Geometria, quam vulgo putatur.' Such language from so great a mathematician invited confusion in weaker minds;† and Leibniz's 'lex continuitatis', 'unde fit, ut in continuis extremum exclusivum tractari possit ut inclusivum'—the principle, so often appealed to by the British mathematicians of the early nineteenth century, that 'what is true up to the limit is true at the limit'—was still more unfortunate. It was nearly 100 years later when Lagrange (referring to an observation of Callet which we shall quote in a moment) remarked that 'les géomètres doivent savoir gré au cit. Callet d'avoir appelé leur attention sur l'espèce de paradoxe que présentent les séries dont il s'agit, et d'avoir cherché à les prémunir contre l'application des raisonnements métaphysiques aux questions qui, n'étant que de pure analyse, ne peuvent être décidées que par les premiers principes et les règles fondamentales du calcul'.

Callet's remark refers to Euler's principle 'summa cujusque seriei ...', which we quoted in § 1.3, and which was the subject of a correspondence between Euler and N. Bernoulli in 1743. Bernoulli had objected that the same series might 'arise' from two different 'expressions' which yielded different values, and Euler had committed himself to the assertion that this could not happen. Writing to Goldbach in 1745, he says 'Darüber hat er zwar kein Exempel gegeben, ich glaube aber gewiss zu sein, dass nimmer eben dieselbe series aus der Evolution zweier wirklich verschiedenen expressionum finitorum entstehen könne'. Callet, forty or fifty years later, observed that $1-1+1-...$ arises, when we put $x = 1$, not only from $(1+x)^{-1} = 1-x+x^2-...$, but also from

$$\frac{1+x+...+x^{m-1}}{1+x+...+x^{n-1}} = \frac{1-x^m}{1-x^n} = 1-x^m+x^n-x^{n+m}+x^{2n}-..., \tag{1.6.3}$$

for any $m < n$, and that Euler's principle might thus be made to give any sum m/n for $1-1+1-...$.

The explanation is fairly obvious (and was given by Lagrange himself). The series (1.6.3), considered as a power series, has gaps: thus when $m = 2$, $n = 3$, it is

$$1+0.x-1.x^2+1.x^3+0.x^4-1.x^5+...;$$

Euler's principle does not assign the sum $\frac{2}{3}$ to $1-1+1-...$ but to $1+0-1+1+0-1+...$; and there is no *a priori* reason for expecting

† Even Euler appealed to metaphysics when he could think of nothing better—'per rationes metaphysicas . . . quibus in analysi acquiescere queamus'.

the two series to have the same sum. And, in fact, Euler's assertion, when properly interpreted, is correct, since a convergent power series has a unique generating function.

It is a mistake to think of Euler as a 'loose' mathematician, though his language may sometimes seem loose to modern ears; and even his language sometimes suggests a point of view far in advance of the general ideas of his time. Thus, in the very passage in which he formulates his principle, he refers to the series (1.1.4). The principle, as we formulated it in § 1.3, does not apply to this series, since $1-1!x+2!x^2-3!x^3+\dots$ is not convergent for any x but 0. Even so, says Euler, 'ich glaube, dass jede series einen bestimmten Wert haben müsse. Um aber allen Schwierigkeiten, welche dagegen gemacht worden, zu begegnen, so sollte dieser Wert nicht mit dem Namen der Summe belegt werden, weil man mit diesem Wort gemeiniglich einen solchen Begriff zu verknüpfen pflegt, als wenn die Summe durch eine wirkliche Summierung herausgebracht würde: welche Idee bei den seriebus divergentibus nicht stattfindet. . . .' This is language which might almost have been used by Cesàro or Borel. And in another place, referring more generally to the controversies excited by the use of divergent series, he suggests that they are largely *verbal*: 'quemadmodum autem iste dissensus realis videatur, tamen neutra pars ab altera ullius erroris argui potest, quoties in analysi hujusmodi serierum usus occurrit: quod gravi argumento esse debet, neutram partem in errore versari, sed totum dissidium in solis verbis esse positum.' Here, as elsewhere, Euler was substantially right. The puzzles of the time about divergent series arose mostly, not from any particular mystery in divergent series as such, but from disinclination to give formal definitions and from the inadequacy of the current theory of functions. It is impossible to state Euler's principle accurately without clear ideas about functions of a complex variable and analytic continuation.

(3) It is essential to remember that Euler was thinking of power series; as soon as we admit other kinds of development, all sorts of difficulties appear. Thus

$$(1-2x)^{-1} = 1+2x+4x^2+8x^3+\dots$$

gives

$$1+2+4+8+\dots = -1;$$

and (1.2.1) gives (the complex) ∞ as the sum of $1+1+1+\dots$. But

$$\frac{2}{e^{2y}-1} = \frac{2}{e^{2y}+1}+\frac{4}{e^{4y}+1}+\frac{8}{e^{8y}+1}+\dots \quad (y>0)\dagger$$

† This is a corollary of $\dfrac{1}{x-1} = \dfrac{1}{x+1}+\dfrac{2}{x^2-1}$.

gives
$$1+2+4+8+\ldots = \infty$$
for $y = 0$, and
$$\zeta(s) = 1^{-s}+2^{-s}+3^{-s}+\ldots \quad (s > 1)$$
gives
$$1+1+1+\ldots = \zeta(0) = -\tfrac{1}{2}$$
for $s = 0$. On the other hand,
$$x+(2x^2-x)+(3x^3-2x^2)+(4x^4-3x^3)+\ldots = 0$$
and
$$x+(3x^2-x)+(7x^4-3x^2)+(15x^8-7x^4)+\ldots = 0$$
for $0 \leqslant x < 1$, and these give
$$1+1+1+\ldots = 0, \qquad 1+2+4+\ldots = 0$$
for $x = 1$.

There are also difficulties, even for power series, with many-valued functions. It is natural to say that
$$\tfrac{2}{1}-\tfrac{4}{2}+\tfrac{8}{3}-\ldots = \log(1+2) = \log 3,$$
since $\log 3$ is the value of $\log(1+x)$ when x moves to 2 in the obvious way. But we might also argue that
$$\tfrac{2}{1}+\tfrac{4}{2}+\tfrac{8}{3}+\ldots = \log(1-2)^{-1} = \log(-1) = (2k+1)\pi i,$$
and here πi and $-\pi i$ seem equally natural values (though either has an air of paradox).

The following example might have puzzled Euler. The series

$$1+\frac{1}{2}\left(\frac{2x}{1+x^2}\right)^2+\frac{1.3}{2.4}\left(\frac{2x}{1+x^2}\right)^4+\ldots \tag{1.6.4}$$

is convergent for small and also for large x, but to different sums, viz.
$$(1+x^2)/(1-x^2), \qquad (x^2+1)/(x^2-1)$$
respectively. If $x = 2i$ we obtain
$$1-\frac{1}{2}\,\frac{16}{9}+\frac{1.3}{2.4}\left(\frac{16}{9}\right)^2-\ldots = \pm\tfrac{3}{5}.$$
Which sign shall we choose?

(4) It is interesting in this connexion to look at a transformation of the geometric series which is due to Goldbach and which may be regarded as an eighteenth-century essay in 'analytic continuation'. We have simplified and generalized Goldbach's actual analysis.

The idea is to transform $1-x+x^2-\ldots$, by formal multiplication by a series of the type
$$1+A_1-A_1+A_2-A_2+\ldots = 1,$$
into a series of negative powers of $y = ax+b$. We write $A_n = \alpha_n y^{-n}$ and arrange the product as

$$\begin{array}{ccccc}
1 & -x & x^2 & -x^3 & \ldots \\
\alpha_1 y^{-1} & -\alpha_1 y^{-1}-\alpha_1 xy^{-1} & \alpha_1 xy^{-1}+\alpha_1 x^2y^{-1} & -\alpha_1 x^2y^{-1}-\alpha_1 x^3y^{-1} & \ldots \\
 & \alpha_2 y^{-2} & -\alpha_2 y^{-2}-\alpha_2 xy^{-2} & \alpha_2 xy^{-2}+\alpha_2 x^2y^{-2} & \ldots \\
 & & \alpha_3 y^{-3} & -\alpha_3 y^{-3}-\alpha_3 xy^{-3} & \ldots \\
 & & & \ldots & \ldots
\end{array}$$

If now we take $\alpha_n = (b-a)^{n-1}(a-y)$, then it will be found† that C_n, the sum of the nth column, is $a(b-a)^{n-1}y^{-n}$, and we obtain

$$1-x+x^2-x^3+\ldots = ay^{-1}+a(b-a)y^{-2}+a(b-a)^2y^{-3}+\ldots,$$

which is the expansion required. The first series is convergent for $|x| < 1$, the second for $|ax+b| > |b-a|$. If, for example, $b > a > 0$, then the second region includes the first. Since both series are convergent, and the transformation valid, for $|x| < 1$, the second series gives the continuation of the first.

(5) Mathematics after Euler moved slowly but steadily towards the orthodoxy ultimately imposed on it by Cauchy, Abel, and their successors, and divergent series were gradually banished from analysis, to reappear only in quite modern times. They had always had their opponents, such as d'Alembert,‡ Laplace,§ and (in his later days) Lagrange: after Cauchy, the opposition seemed definitely to have won.

The analysts who used divergent series most, after Euler, were Fourier and Poisson (who was almost Cauchy's contemporary). We shall see specimens of their work in Chs. II and XIII. The most important for us here is Poisson, since he so nearly formulated definition (2) of § 1.3. Poisson, in effect, defines the sum of the trigonometrical series

$$\tfrac{1}{2}a_0+\sum(a_n\cos n\theta+b_n\sin n\theta)$$

as the limit when $r \to 1$ of the associated power series

$$\tfrac{1}{2}a_0+\sum(a_n\cos n\theta+b_n\sin n\theta)r^n.$$

Thus, speaking of the series (1.1.9), he says 'cette série n'est ni convergente ni divergente‖ et ce n'est qu'en la considérant ainsi que nous le faisons comme la limite d'une série convergente, qu'elle peut avoir une valeur déterminée. . . . Nous admettrons avec Euler que les sommes de ces séries considérées en elles-même n'ont pas de valeurs déterminées; mais nous ajouterons que chacune d'elles a une valeur unique et qu'on peut les employer dans l'analyse, lorsqu'on les regarde comme les limites des séries convergentes, c'est à dire quand on suppose implicitement leurs termes successifs multipliés par les puissances d'une fraction infiniment peu différente de l'unité.' This is practically the 'A' definition, but we must not exaggerate the clarity of Poisson's views. His ideas

† By induction from $C_n = -xC_{n-1}-\alpha_{n-1}y^{-n+1}+\alpha_n y^{-n}$.

‡ 'Pour moi j'avoue que tous les raisonnements et les calculs fondés sur des séries qui ne sont pas convergentes . . . me paraîtront toujours très suspects, même quand les résultats de ces raisonnements s'accorderaient avec des vérités connues d'ailleurs.'

§ 'Je mets encore au rang des illusions l'application que Leibniz et Dan. Bernoulli ont faite du calcul des probabilités . . . (to the summation of such series as $1-1+1-\ldots$).'

‖ He means, of course, 'properly divergent' to ∞ or $-\infty$.

about repeated limits are often by no means clear: thus he writes Fourier's theorem as

$$f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt + \frac{1}{\pi}\int_{-\pi}^{\pi}\left\{\sum_{1}^{\infty}\cos n(t-x)\right\}f(t)\,dt,$$

when, of course, he *means*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt + \frac{1}{\pi}\sum_{1}^{\infty}\int_{-\pi}^{\pi}\cos n(t-x)f(t)\,dt.$$

1.7. A note on the British analysts of the early nineteenth century. We end this chapter with a few remarks about British work on these subjects during the years 1840–50, which has been analysed very carefully by Burkhardt in the article from which we have quoted. It was a long time before the writings of the great continental analysts were understood in England, and these British writings show a singular and often entertaining mixture of occasional shrewdness and fundamental incompetence.

(1) The dominant school was that of the Cambridge 'symbolists', Woodhouse, Peacock, D. F. Gregory, and others. They represented what may be described as the '$f(D)$' school of analysis. They started from 'algebra', and had something of the spirit, though nothing of the accuracy, of the modern abstract algebraists. They dealt in 'general symbols', on which operations were to be performed in accordance with certain laws: 'the symbols are unlimited, both in value and in representation; the operations upon them, whatever they may be, are possible in all cases; . . .' But the foundations of their symbolism were both inelastic and inaccurate. They insisted on a parallelism between 'arithmetical' and 'general' algebra so rigid that, if it could be maintained, it would effectively destroy the generality; and they never seem to have realized fully that a formula true with one interpretation of its symbols is quite likely to be false with another. They were also very much at the mercy of catchwords like 'what is true up to the limit . . .', and it is not surprising that their permanent contribution to analysis should have been negligible.

Occasionally, however, they arrived at formulae which are still worth examining. Thus Gregory's formulae

(1.7.1) $\sum_{-\infty}^{\infty}\phi(x+n) = 0,$ (1.7.2) $\sum_{-\infty}^{\infty}(-1)^n\phi(x+n) = 0,$

(1.7.3) $$\sum_{1}^{\infty}\frac{(-1)^{n-1}}{n}\{\phi(x+n)-\phi(x-n)\} = \phi'(x),$$

are true, or true with modifications, when interpreted properly, for interesting classes of functions.

(2) There is one volume of the *Transactions of the Cambridge Philosophical Society* (vol. 8, published in 1849 and covering the period 1844–9) which contains a very singular mixture of analytical papers and gives a particularly good picture of the British analysis of the time. It contains Stokes's famous paper 'On the critical values of the sums of periodic series', in which 'uniform convergence'

appears first in print; papers by S. Earnshaw and J. R. Young which are little more than nonsense; and a long and interesting paper by de Morgan on divergent series, a remarkable mixture of acuteness and confusion.

De Morgan, as Burkhardt recognizes, was no 'blosser Algorithmiker' like Peacock. He was a prolific and ingenious writer, both on logic and on mathematics; he invented the 'logarithmic scale' of convergence criteria; and his *Differential and integral calculus*, which is the best of the early English text-books on the calculus, contains much that is still interesting to read and difficult to find in any other book. In this paper he attempts a reasoned statement of his attitude to divergent series, 'the only subject yet remaining, of an elementary character, on which serious schism exists among mathematicians as to absolute correctness or incorrectness of results'. He talks much excellent sense, but the habits of the time are too strong for him: logician though he is, he cannot, or will not, give definitions.

'The moderns', he says, 'seem to me to have made a similar confusion in regard to their rejection of divergent series; meaning sometimes that they cannot safely be used under existing ideas as to their meaning and origin, sometimes that the mere idea of anyone applying them at all, under any circumstances, is an absurdity. We must admit that many series are such as we cannot safely use, except as means of discovery, the results of which are to be subsequently verified. . . . But to say that what we cannot use no others ever can . . . seems to me a departure from all rules of prudence. . . .' Would analysis ever have developed as it has done if Euler and others had refused to use $\sqrt{(-1)}$?

He refuses to distinguish between different types of divergent series: if some are to be used, all must be. 'I do not argue with those who reject everything that is not within the province of arithmetic, but only with those who abandon the use of infinitely divergent series and yet appear to employ finitely divergent series with confidence. Such appears to be the practice, both at home and abroad. They seem perfectly reconciled to $1-1+1-... = \frac{1}{2}$, but cannot admit $1+2+4+... = -1$.' It is very odd that it should never have occurred to him that there might be interpretations (for example, Poisson's) which apply in the one case and not in the other.

Later, when he recurs to this point, he is a little inconsistent. There are cases in which $1+2+4+...$ seems to represent -1, others in which it seems to represent ∞:† thus the limit of

$$1+2x+4x^4+...+2^n x^{n^2}+...,$$

as $x \to 1$, is ∞ (a well-chosen example). This he can tolerate, but 'let it come out anything but -1 or ∞, and as a result of any process which does not involve *integration performed on a divergent series* . . . and I shall then be obliged to admit that divergent series must be abandoned'.‡ There is *something* in his view: -1 is a root of $z = 1+2z$, and there is a sense in which ∞ can be said to be one also, while 0 or 1 certainly cannot. We found 0 in (3) of § 1.6, but de Morgan would certainly have felt that the example was unfair, and would not have been altogether wrong. It is true that -1 and ∞ are the only 'natural' sums.

Similarly with $1-1+1-...$: it would be fatal if this came out to be anything

† See § 1.6 (3).

‡ The emphasis on integration is odd, but de Morgan seems to have regarded integration as an 'essentially arithmetic' process liable to destroy any more 'symbolic' reasoning.

other than $\frac{1}{2}$. 'The whole fabric of periodic series and integrals . . . would fall instantly if it were shown to be possible that $1-1+1-\ldots$ might be one quantity as a limiting form of $A_0-A_1+A_2-\ldots$ and another as a limiting form of $B_0-B_1+B_2-\ldots$'; and here there is some quite mistaken criticism of Poisson. De Morgan implies that to define $\sum a_n$ as $\lim \sum a_n x^n$ is to assume that 'what is true up to the limit is true at the limit'; whereas it is just this distinction which is seized upon by, and embodied in, the definition.

He gives curious examples of paradoxes resulting from integration. Here and elsewhere he shows a good deal of formal ingenuity, but other paradoxes rest merely on confusion about many-valued functions. He forgets that the integral of x^{-1} is $\log|x|$, not $\log x$, when x is negative, and concludes that

$$\int_0^{2\pi} \tan x\, dx = 2\pi i$$

and that '$\tan^2 x$ has -1 for its mean value'—a conclusion which he tries to reinforce on other grounds. There is also some discussion of the formulae (1.7.1)–(1.7.3), and of alternating asymptotic series of the Euler-Maclaurin type. 'When an alternating series is convergent, and a certain number of its terms are taken . . . the first term neglected is a superior limit to the error of approximation. . . .† This very useful property was observed to belong to large classes of alternating series, when finitely or even infinitely divergent: I do not remember that anyone has *denied* that it is universally true. . . .' De Morgan shows by examples that it is not, but without making any substantial contribution to the subject. Indeed these supplementary discussions merely confirm the impression left by the earlier sections of the paper, of astonishment that so acute a reasoner should be able to say so much that is interesting and yet to miss the essential points so completely.

(3) It is only fair to quote a few instances of British analysts who got nearer to actuality. F. W. Newman protested against the dogma 'what is true . . .' and pointed out that, in the case of the trigonometrical series

$$\cos x - \tfrac{1}{3}\cos 3x + \tfrac{1}{5}\cos 5x - \ldots,$$

it is plainly false. His analysis is unsatisfactory, but he makes his point substantially, and his paper is interesting because it led Wilbraham, a little later, to the discovery of what is now called the 'Gibbs phenomenon'. Stokes, in his famous paper mentioned already, remarked that 'of course we may employ a divergent series merely as an abbreviated way of expressing the limit of the sum of a convergent series', and observed that it did not seem possible 'to invent a series so rapidly divergent that it shall not be possible to find a convergent series which shall have, for the limits of its first n terms, the first n terms of the divergent series'.‡ Finally, Homersham Cox, referring to the 'equivalence' of the symbolists, used language entirely modern in spirit: 'it is said that the symbol "=" here designates symbolical equivalence. The truth of this assertion depends on the definition of this phrase, and without doubt many arbitrary definitions might be given, in accordance with which the binomial theorem might be considered to hold for divergent series.'

† Of course this is not true without reservation.

‡ Consider $\sum \phi(n)x^{\phi(n)}$, where $\phi(n) \to \infty$ rapidly.

NOTES ON CHAPTER I

§ 1.1. Many writers, particularly in England, have used 'divergent' in narrower senses. Thus Bromwich, Hardy, and Hobson, in their text-books, call $\sum a_n$ divergent only when $s_n \to \infty$ or $s_n \to -\infty$, describing other non-convergent series as 'oscillatory'. In his first edition Hobson had called $\sum a_n$ divergent if $|s_n| \to \infty$: thus $1-2+3-\ldots$ was divergent.

The narrower use of 'divergent' has its advantages in elementary teaching, but the wider use is almost necessary here. The 'theory of divergent series' is essentially a theory of oscillatory series, theorems about series which diverge 'properly' to ∞ or $-\infty$ being usually of the same type as those about convergent series. See, for example, § 3.6, and the remarks in Hobson, 2, 4.

Cauchy's *Analyse algébrique* (Paris, 1821) was the first standard treatise on analysis written in a genuinely modern spirit. A good deal of his work on the foundations is to be found, sometimes even in a sharper form, in a series of memoirs published by Bolzano in Prag in 1817. See Stolz, *MA*, 18 (1881), 255–79.

§ 1.2. For justification of the results in (1)–(3), (5), and (6) see Appendix I. As regards (4), if

$$f(x) = \int_0^a \cos xt\, \chi(t)\, dt,$$

where $0 < a \leqslant 1$, $\chi(t)$ is any integrable function, and $-\pi \leqslant \theta \leqslant \pi$, then

$$\sum_1^\infty (-1)^{n-1}\frac{f(n\theta)}{n^2} = \int_0^a \left\{\sum_1^\infty (-1)^{n-1}\frac{\cos n\theta t}{n^2}\right\}\chi(t)\, dt = \int_0^a \left(\frac{\pi^2}{12}-\frac{\theta^2 t^2}{4}\right)\chi(t)\, dt,$$

which is $\frac{1}{12}\pi^2 f(0)+\frac{1}{4}\theta^2 f''(0)$, in agreement with (1.2.24). Many other formulae of the same kind may be proved similarly. The limitations $a \leqslant 1$ and $|\theta| \leqslant \pi$ are essential.

For the Bernoullian and Eulerian numbers, and the Bernoullian functions, see Bromwich, 297 et seq., 370, and Chapter XIII.

The series (1.2.18) seems to have first been summed by integration by Euler, *Novi Commentarii Acad. Petropolitanae*, 5 (1760), 203. [*Opera* (I), 14, 542–84. He gives another method in *Opera* (I), 15, 435–97.]

§ 1.3. More detailed information about the early work of Bernoulli and others on divergent series will be found in Reiff, *Geschichte der unendlichen Reihen* (Tübingen, 1889), in a paper by Burkhardt in *MA*, 70 (1911), 169–206, and in Burkhardt's article 'Trigonometrische Reihen und Integrale' in the *Enzykl. d. Math. Wiss.* (IIA 12). Reiff's book is useful but uninspiring and not always accurate. Burkhardt's writings are much more interesting, and contain a mass of curious information difficult to find elsewhere. The historical discussions here and in §§ 1.6–7 are based mainly on these sources.

Hutton, *Tracts on math. and philosophical subjects* (London, 1812), gave what is in effect the following definition of the limit of a divergent sequence (s_n). Define $s_n^{(k)}$ for $k = 1, 2,\ldots$ by

$$s_n^{(k)} = \tfrac{1}{2}s_{n-1}^{(k-1)}+\tfrac{1}{2}s_n^{(k-1)} \qquad (n \geqslant 0),$$

with

$$s_{-1}^{(k)} = 0 \quad (k \geqslant 0), \qquad s_n^{(0)} = s_n \quad (n \geqslant 0).$$

Thus $s_n^{(1)}$ and $s_n^{(2)}$ are

$$\tfrac{1}{2}s_0,\ \tfrac{1}{2}s_0+\tfrac{1}{2}s_1,\ \tfrac{1}{2}s_1+\tfrac{1}{2}s_2,\dots,\ \tfrac{1}{2}s_{n-1}+\tfrac{1}{2}s_n,\dots,$$

$$\tfrac{1}{4}s_0,\ \tfrac{1}{2}s_0+\tfrac{1}{4}s_1,\ \tfrac{1}{4}s_0+\tfrac{1}{2}s_1+\tfrac{1}{4}s_2,\dots,\ \tfrac{1}{4}s_{n-2}+\tfrac{1}{2}s_{n-1}+\tfrac{1}{4}s_n,\dots.$$

Then $$s_n \to s\ (\mathrm{Hu}, k) \quad \text{means} \quad s_n^{(k)} \to s.$$

It is easily verified that

$$1-1+1-\dots = \tfrac{1}{2}\ (\mathrm{Hu}, 1), \qquad 1-2+3-\dots = \tfrac{1}{4}\ (\mathrm{Hu}, 2);$$

and we can show, from the general theorems of Ch. III, or those about Nörlund means in Ch. IV, that any series summable (Hu, k) is summable, to the same sum, by the corresponding Cesàro mean.

For examples of the use of Euler's transformation in numerical computation see Bromwich, 62–6.

§ 1.4. For Cauchy's theorem see Hardy, 167, or Bromwich, 414. It is a case of Theorem 44. Abel's theorem is included, for example, in Theorems 27 and 55.

§ 1.5. For (1.5.10) and (1.5.11) see Appendix I, § 4, where some errors in a paper in *TCPS*, 21 (1908), 1–48, are rectified.

§ 1.6. The first criterion for convergence formulated explicitly seems to have been Leibniz's familiar criterion for the convergence of an alternating series $a_0-a_1+a_2-\dots$ with positive decreasing a_n.

(3) The series (1.6.4) converges in two regions bounded by the circles $u^2+(v\pm 1)^2 = 2$, where $u+iv = x$, the lune inside both and the infinite region outside both; and diverges in the two remaining lunes. The point $2i$ is in the upper of these last two lunes. The series represents a single two-valued function of $z = 2x/(1+x^2)$, but two different one-valued functions of x.

(4) For Goldbach's actual statement of the transformation see M. Cantor, *Vorlesungen über Geschichte der Math.*, 3, ed. 2 (Leipzig, 1901), 641. The account in Reiff, 89, is incorrect.

§ 1.7 (1). A reader acquainted with the elements of the theory of Fourier series will easily verify the truth of (1.7.1)–(1.7.3) for $\phi(x)$ defined by appropriate trigonometrical integrals.

(2) Burkhardt analyses the papers of Earnshaw and Young with more care than they deserve. He also says a good deal about minor German work of the same period, but this is on the whole less interesting.

(3) The papers of Newman, Wilbraham, and Homersham Cox appeared in the *Cambridge and Dublin Math. Journal*, 3 (1848), 108 and 198, and 7 (1852), 98. F. W. Newman, Professor of Mathematics in University College, London, was a brother of Cardinal J. H. Newman.

II

SOME HISTORICAL EXAMPLES

2.1. Introduction. In this chapter we give the examples of the work of Euler and others which were promised in § 1.6, starting in each case from a passage in the original writings of the analyst in question. The subject-matter of these passages is still important, so that they have more than an historical interest; and we shall therefore analyse them in some detail, and add the explanations needed to show their connexion with more modern work.

A. *Euler and the functional equation of Riemann's zeta-function*

2.2. The functional equations for $\zeta(s)$, $\eta(s)$, and $L(s)$. The Riemann ζ-function $\zeta(s)$, defined by the series

(2.2.1) $$\zeta(s) = 1^{-s}+2^{-s}+3^{-s}+\dots$$

when $s = \sigma+it$ and $\sigma > 1$, is a one-valued analytic function of s, regular all over the plane except for a simple pole at $s = 1$. It satisfies the functional equation

(2.2.2) $$\zeta(1-s) = 2(2\pi)^{-s}\cos\tfrac{1}{2}s\pi\,\Gamma(s)\zeta(s).$$

Near $s = 1$,

(2.2.3) $$\zeta(s) = \frac{1}{s-1}+\gamma+\dots,$$

where γ is Euler's constant.

The functions $\eta(s)$ and $L(s)$, defined for $\sigma > 0$ by

(2.2.4) $\eta(s) = 1^{-s}-2^{-s}+3^{-s}-\dots,$ (2.2.5) $L(s) = 1^{-s}-3^{-s}+5^{-s}-\dots,$

are integral functions of s; $\eta(s) = (1-2^{1-s})\zeta(s)$, but $L(s)$ is an independent transcendent. They satisfy

(2.2.6) $$(2^{s-1}-1)\eta(1-s) = -(2^s-1)\pi^{-s}\cos\tfrac{1}{2}s\pi\,\Gamma(s)\eta(s),$$

(2.2.7) $$L(1-s) = 2^s\pi^{-s}\sin\tfrac{1}{2}s\pi\,\Gamma(s)L(s).$$

These results have usually been attributed to Riemann, Malmstén, and Schlömilch. It was comparatively recently that it was observed, first by Cahen and then by Landau, that both (2.2.6), which is equivalent to (2.2.2), and (2.2.7) stand in a paper of Euler's written in 1749, over 100 years before Riemann. Euler does not consider complex values of s, and does not profess to have *proved* the equations even for real s. He states them, and verifies them in such a number of cases as 'ne plus laisser aucun doute sur la vérité de notre conjecture'. Incidentally his verifications throw much light on his views about divergent series.

2.3. Euler's verification. Euler states (2.2.6) in the form

$$\frac{1-2^{s-1}+3^{s-1}-\dots}{1-2^{-s}+3^{-s}-\dots} = -\frac{(s-1)!(2^s-1)}{(2^{s-1}-1)\pi^s}\cos\tfrac{1}{2}s\pi, \tag{2.3.1}$$

and proceeds to verify this equation (*a*) for all integral s and (*b*) for $s=\frac{1}{2}$ and $s=\frac{3}{2}$. It will be observed that $s=\frac{1}{2}$ is the only one of these values of s for which both series are convergent.

He needs the formulae

$$1-2^{-2k}+3^{-2k}-\dots = \frac{2^{2k-1}-1}{(2k)!}\pi^{2k}B_k, \tag{2.3.2}$$

$$1-1+1-\dots = \tfrac{1}{2}\ (A), \tag{2.3.3}$$

$$1-2^{2k}+3^{2k}-\dots = 0\ (A), \tag{2.3.4}$$

$$1-2^{2k-1}+3^{2k-1}-\dots = (-1)^{k-1}\frac{2^{2k}-1}{2k}B_k\ (A). \tag{2.3.5}$$

Here k is a positive integer. Of these formulae (2.3.2) is familiar, and the others, apart from the (A), are (1.2.7), (1.2.12), and (1.2.13).

It is important to observe that, here at any rate, Euler is quite explicit about his use of divergent series: the series are to be summed by the A definition of § 1.3(2). It is easy to verify their truth in this sense. For from

$$e^{-y}-e^{-2y}+e^{-3y}-\dots = (e^y+1)^{-1} \quad (y>0) \tag{2.3.6}$$

it follows that

$$1^m e^{-y}-2^m e^{-2y}+3^m e^{-3y}-\dots = (-1)^m\left(\frac{d}{dy}\right)^m\frac{1}{e^y+1} \tag{2.3.7}$$

for $m=0, 1, 2,\dots$. Now

$$\frac{1}{e^y+1} = \tfrac{1}{2}-\tfrac{1}{2}\tanh\tfrac{1}{2}y = \tfrac{1}{2}-\sum_{1}^{\infty}(-1)^{k-1}\frac{2^{2k}-1}{(2k)!}B_k y^{2k-1},$$

so that the limit of (2.3.7) is

$$\tfrac{1}{2}\ (m=0),\quad 0\ (m=2k>0),\quad (-1)^{k-1}\frac{2^{2k}-1}{2k}B_k\ (m=2k-1>0).$$

It follows that the series (2.3.6) and (2.3.7) have the limits required when $y\to 0$ or $x=e^{-y}\to 1$.

This proves (2.3.3)–(2.3.5), and also shows that the series are summable $(\mathfrak{E})$ to the same sum: Euler might equally well have used this definition. We can naturally prove the truth of (1.2.8)–(1.2.11), (1.2.14), and (1.2.16), in the same senses, in the same way.

From (2.3.4) $\eta(1-s) = 0 \quad (s = 3, 5, 7,...);$

and from (2.3.2) and (2.3.5)

$$\frac{\eta(1-s)}{\eta(s)} = \frac{(-1)^{\frac{1}{2}s-1}(s-1)!(2^s-1)}{(2^{s-1}-1)\pi^s} \quad (s = 2, 4, 6,...).$$

If we observe that $\cos\frac{1}{2}s\pi$ is 0 when s is odd and $(-1)^{\frac{1}{2}s}$ when s is even, then these two formulae verify (2.3.1) for $s = 2, 3, 4, 5,...$.

Secondly, if we take $s = 1$, and interpret $\cos\frac{1}{2}s\pi/(2^{s-1}-1)$, for $s = 1$, as its limit when $s \to 1$, i.e. as $-\pi/(2\log 2)$, then (2.3.1) becomes

$$\frac{1-1+1-...}{1-\frac{1}{2}+\frac{1}{3}-...} = \frac{1}{2\log 2},$$

in agreement with (2.3.3).

Thirdly, if we write $(s-1)!(2^s-1) = s!(2^s-1)/s$, and interpret this as $1.\log 2$ for $s = 0$, then (2.3.1), for $s = 0$, becomes

$$\frac{1-\frac{1}{2}+\frac{1}{3}-...}{1-1+1-...} = 2\log 2,$$

again in agreement with (2.3.3).

Fourthly, replacing $(s-1)!$ by $\Gamma(s)$, and using

$$\Gamma(s)\cos\tfrac{1}{2}s\pi = \frac{\pi}{2\Gamma(1-s)\cos\frac{1}{2}(1-s)\pi},$$

we find that the truth of (2.3.1) for general $s > 1$ implies its truth for $s < 0$. We may then regard the formula as verified for all integral s.

Fifthly, if $s = \frac{1}{2}$, and we interpret $(-\frac{1}{2})!$ as $\Gamma(\frac{1}{2})$, then

$$-\frac{(s-1)!(2^s-1)}{(2^{s-1}-1)\pi^s}\cos\tfrac{1}{2}s\pi = -\frac{\Gamma(\frac{1}{2})(2^{\frac{1}{2}}-1)}{(2^{-\frac{1}{2}}-1)\sqrt{\pi}\sqrt{2}} = 1,$$

so that (2.3.1) is true for $s = \frac{1}{2}$. This completes Euler's programme except for the value $s = \frac{3}{2}$. For this he has only a numerical verification. He sums the divergent series with the help of the Euler-Maclaurin sum formula, and finds the value ·380129.... This gives ·496774 for the value of the left-hand side of (2.3.1), in agreement with the right to 5 figures. 'Notre conjecture est portée au plus haut degré de certitude, qu'il ne reste plus même aucun doute sur les cas où l'on met pour l'exposant s des fractions.'

As Landau remarks, Euler's computation of $1-\sqrt{2}+\sqrt{3}-...$ may easily be transformed into a rigorous determination of its Abel sum. It is worth observing that the sum may also be calculated by Euler's transformation of § 1.3(4). In this case $a_n = (-1)^n\sqrt{(n+1)}$ and, calculating the successive differences of $\sqrt{(n+1)}$, we find

$$b_0 = 1,\ b_1 = -·4142,\ b_2 = -·0964,\ b_3 = -·0465,\ b_4 = -·0285,\ b_5 = -·0197,$$

so that Euler's series is

$$\frac{1}{2}-\frac{\cdot 4142}{4}-\frac{\cdot 0964}{8}-\frac{\cdot 0465}{16}-\frac{\cdot 0285}{32}-\frac{\cdot 0197}{64}-\ldots,$$

which is ·380 to 3 figures. Our calculation is of course much rougher than Euler's.

Euler does not discuss (2.2.7) in the same detail, but implies that he has made similar verifications. He ends by remarking that 'cette dernière conjecture renferme une expression plus simple que la précédente; donc, puisqu'elle est également certaine, il y a à espérer qu'on travaillera avec plus de succès à en chercher une démonstration parfaite, qui ne manquera pas de répandre beaucoup de lumière sur quantité d'autres recherches de cette nature'.

B. *Euler and the series* $1-1!x+2!x^2-\ldots$

2.4. Summation of the series. The series

$$f(x) = 1-1!x+2!x^2-3!x^3+\ldots, \tag{2.4.1}$$

which reduces to (1.1.4) for $x = 1$, is not convergent for any x except $x = 0$, or summable by any of the methods of § 1.3. For example, when $x = 1$, the series (1.3.5) diverges almost as rapidly as the original series.

Euler, however, succeeded in summing the series as follows. If we suppose x positive and write, formally,

$$\phi(x) = xf(x) = x-1!x^2+2!x^3-\ldots,$$

then term-by-term differentiation gives

$$x^2\phi'(x)+\phi(x) = x^2(1!-2!x+3!x^2-\ldots)+x-1!x^2+\ldots = x. \tag{2.4.2}$$

This differential equation has the integrating factor $x^{-2}e^{-1/x}$, and

$$\phi(x) = e^{1/x}\int_0^x \frac{e^{-1/t}}{t}\,dt \tag{2.4.3}$$

is a solution which vanishes with x.†

If we make the substitution $t = x/(1+xw)$, we obtain

$$f(x) = \frac{\phi(x)}{x} = \int_0^\infty \frac{e^{-w}}{1+xw}\,dw; \tag{2.4.4}$$

and it is natural to attribute this sum to the series (2.4.1), the more so because we come back to the series by expanding $(1+xw)^{-1}$ in powers of xw and integrating formally term by term.

† It is easily verified by partial integration that $\phi(x) = O(x)$ for small x.

Since $x > 0$, we have also, from (2.4.3),

$$f(x) = \frac{1}{x}e^{1/x}\int_0^x \frac{e^{-1/t}}{t}\,dt = \frac{1}{x}e^{1/x}\int_{1/x}^{\infty} \frac{e^{-u}}{u}\,du = -\frac{1}{x}e^{1/x}\operatorname{li}(e^{-1/x}), \tag{2.4.5}$$

where $\operatorname{li} v$, the 'logarithm-integral' of v, is defined for $0 < v < 1$ by

$$\operatorname{li} v = \int_0^v \frac{dt}{\log t} = -\int_{\log(1/v)}^{\infty} \frac{e^{-u}}{u}\,du.$$

Then

$$-\operatorname{li}(e^{-y}) = \int_y^{\infty} \frac{e^{-u}}{u}\,du = \int_1^{\infty} \frac{e^{-u}}{u}\,du - \int_0^1 \frac{1-e^{-u}}{u}\,du - \int_1^y \frac{du}{u} + \int_0^y \frac{1-e^{-u}}{u}\,du$$

$$= -\gamma - \log y + y - \frac{y^2}{2.2!} + \frac{y^3}{3.3!} - \dots,$$

and it follows from (2.4.5) that

$$f(x) = -\frac{1}{x}e^{1/x}\log\frac{1}{x} + S\left(\frac{1}{x}\right), \tag{2.4.6}$$

where

$$S(y) = -ye^{y}\left(\gamma - y + \frac{y^2}{2.2!} - \frac{y^3}{3.3!} + \dots\right) \tag{2.4.7}$$

is an integral function of y. These equations give the analytic continuation of $f(x)$ all over the plane. It is a many-valued function with an infinity of branches differing by integral multiples of $2\pi i x^{-1}e^{1/x}$, and has one branch which tends to 1 when $x \to 0$ through positive values.

If we take $x = 1$, we obtain the equation

$$1 - 1! + 2! - 3! + \dots = -e\left(\gamma - 1 + \frac{1}{2.2!} - \frac{1}{3.3!} + \dots\right). \tag{2.4.8}$$

2.5. The asymptotic nature of the series. If $x = re^{i\theta}$, where $-\pi < \theta < \pi$, then

$$f(x) = \int_0^{\infty} \frac{e^{-w}}{1+xw}\,dw = \int_0^{\infty} e^{-w}\{1 - xw + x^2w^2 - \dots + (-1)^n x^n w^n\}\,dw + \tag{2.5.1}$$

$$+(-1)^{n+1}x^{n+1}\int_0^{\infty} \frac{e^{-w}w^{n+1}}{1+xw}\,dw = 1 - 1!\,x + 2!\,x^2 - \dots + (-1)^n n!\,x^n + R_n(x),$$

say. Now $|1+xw| = \sqrt{(1+2rw\cos\theta+r^2w^2)}$ has the minimum 1 if $\cos\theta \geqslant 0$, and the minimum $|\sin\theta|$ if $\cos\theta \leqslant 0$. Hence $|R_n(x)|$ does not exceed $(n+1)!\,r^{n+1}$ if $|\theta| \leqslant \frac{1}{2}\pi$, or $(n+1)!\,r^{n+1}|\operatorname{cosec}\theta|$ if $\frac{1}{2}\pi \leqslant \theta < \pi$,

and is $O(r^{n+1})$ uniformly in the angle $-\pi+\delta \leqslant \theta \leqslant \pi-\delta$, for any positive δ. In particular

$$f(x) = 1-1!\,x+2!\,x^2-...+(-1)^n n!\,x^n+O(x^{n+1}) \tag{2.5.2}$$

for small positive x and given n.

A series $$a_0+a_1x+a_2x^2+...$$

is said to be an *asymptotic series for* $f(x)$, near $x=0$, if

$$f(x) = a_0+a_1x+...+a_nx^n+O(x^{n+1}) \tag{2.5.3}$$

for each n and small x. We are interested here primarily in positive x, but the definition applies to complex x also; thus (2.4.1) is an asymptotic series for our $f(x)$ in any angle $-\pi+\delta \leqslant \theta \leqslant \pi-\delta$, i.e. in any angle issuing from the origin and omitting the negative real axis. There is therefore one sense at any rate in which the series 'represents' $f(x)$.

The definition of an asymptotic series is interesting only when the series is divergent. If $f(x)$ is regular at the origin, then its Taylor's series $\sum a_n x^n$ is convergent for small x and satisfies (2.5.3); but in this case there is no novelty in the idea. Divergent asymptotic series occur in the works of most of the older analysts, but the first mathematicians to make a systematic study of them were Poincaré and Stieltjes, and the first general theory is contained in a famous memoir of Poincaré on differential equations.

There will usually be an infinity of different functions represented asymptotically by the same series $\sum a_n x^n$. Thus if $g(x) = e^{-a/x}$, where a is positive, then $x^{-n-1}g(x) \rightarrow 0$ for every n, uniformly in any angle $-\frac{1}{2}\pi+\delta \leqslant \theta \leqslant \frac{1}{2}\pi-\delta$ (and in particular for positive x); so that, for example, the series (2.4.1) gives an asymptotic representation of each of the functions $f(x)+Cg(x)$. To say that a series is an asymptotic series for $f(x)$ is not to 'define its sum' in the sense of § 1.3. There are 'uniqueness theorems' for asymptotic series, due to Watson, F. Nevanlinna, and Carleman; but these depend upon the knowledge of exact bounds for the error terms such as the $R_n(x)$ of (2.5.1), valid for all n and all x of an appropriate region.

We shall often use the phrase 'asymptotic series' in a slightly extended sense, saying that $\sum a_n x^{n+\alpha}$ is an asymptotic series for $f(x)$ if $\sum a_n x^n$ is an asymptotic series for $x^{-\alpha}f(x)$, and we shall sometimes express this by writing

$$f(x) \sim \sum a_n x^{n+\alpha}. \tag{2.5.4}$$

2.6. Numerical computations. Euler calculated a numerical value for the sum of (1.1.4) in various ways. First, we may use (2.4.8). Secondly, we may use (2.4.5), calculating the integral, for $x=1$, by numerical quadrature. These methods give

$$s = 1-1!+2!-3!+... = \cdot 5963.... \tag{2.6.1}$$

There is a more remarkable, though less precise calculation (also due

to Euler) in Lacroix's treatise. Lacroix writes S for $1-s = 1!-2!+\ldots$, and transforms S as in § 1.3(4), obtaining

$$S = \tfrac{1}{2}-\tfrac{1}{4}+\tfrac{3}{8}-\tfrac{11}{16}+\tfrac{53}{32}-\tfrac{309}{64}+\ldots,$$

a series which diverges a little less rapidly than the original series. He then writes $S = \frac{1}{2}-\frac{1}{4}+S'$, and a repetition of the transformation on S' gives

$$S' = \frac{3}{2^4}-\frac{5}{2^6}+\frac{21}{2^8}-\frac{99}{2^{10}}+\frac{615}{2^{12}}-\ldots.$$

Finally, he writes $S' = 3.2^{-4}-5.2^{-6}+S''$, and a third transformation on S'' gives

$$S'' = \frac{21}{2^9}-\frac{15}{2^{12}}+\frac{159}{2^{15}}-\frac{429}{2^{18}}+\frac{5241}{2^{21}}-\ldots.$$

Eight terms of this series lead to the values ·4008 and ·5992 for S and s, correct to two figures. It seems at first very remarkable that we should get so good a result, since all of the series used are divergent (and in the end nearly as rapidly as s). We shall see later (p. 196) why the method should be so successful.

C. *Fourier and Fourier's theorem*

2.7. Fourier's theorem. By 'Fourier's theorem' we mean here the theorem that, if $f(x)$ belongs to an appropriate class of functions, and is 'representable' by a trigonometrical series

(2.7.1) $$\tfrac{1}{2}a_0+\sum_1^{\infty}(a_n\cos nx+b_n\sin nx),$$

in the sense that the series converges to $f(x)$ in the open interval $(-\pi,\pi)$, then

(2.7.2) $$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx, \qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx.$$

Thus the theorem asserts that, if a trigonometrical series converges to $f(x)$ for $-\pi < x < \pi$, then it is necessarily the 'Fourier series' of $f(x)$.

The formulae (2.7.2) are older than Fourier. Thus Burkhardt, in his article in the *Enzyklopädie*, traces the formula for a_n back to Clairaut (1757). They were familiar to Euler, who gave the ordinary deduction of them, by term-by-term integration, in 1777.

It is to be observed that 'Fourier's theorem', as we have stated it, is a 'uniqueness' theorem, and is true or false according to the class of functions considered and the sense of 'representation'. Thus it is true, after du Bois-Reymond and de la Vallée-Poussin, when $f(x)$ is finite

and integrable and representation implies ordinary convergence. If we assume only that the series is summable by one of the standard methods of the theory of divergent series, then the theorem may be false, even when $f(x)$ is always 0. Thus

$$\sin x+2\sin 2x+3\sin 3x+\ldots$$

is summable (A) to 0 for all x, but is obviously not the Fourier series of 0. In any case the theorem is a sophisticated one, which it would have been quite impossible for Fourier to prove strictly: the simplest case of it, in which $f(x)$ is 0 and representation implies convergence, was first proved by Cantor in 1870.

There is a remarkable passage in Fourier's *Théorie de la chaleur* in which he attempts to prove a special case of the theorem. Let us suppose that $f(x)$ is an odd analytic function regular for $|x| \leqslant \pi$, so that

$$f(x) = \sum_{0}^{\infty} \frac{f^{(2k+1)}(0)}{(2k+1)!} x^{2k+1} \tag{2.7.3}$$

for $|x| \leqslant \pi$; and that

$$f(x) = \sum_{1}^{\infty} b_n \sin nx \tag{2.7.4}$$

for $-\pi < x < \pi$, the series being convergent in the classical sense.† These are in effect Fourier's assumptions; and his object is to prove that b_n is given by the second formula (2.7.2).

2.8. Fourier's first formula for the coefficients. The 'natural' method for the proof of (2.7.2) is that of term-by-term integration, which had already been followed by Euler, and would have led Fourier at once to a proof satisfactory according to the canons of the time. Fourier, who does not seem to have known Euler's work, follows a quite different and very surprising course (though he refers to the proof by integration later). He replaces every sine in (2.7.4) by its Taylor's series, and equates the coefficients of powers of x to those in (2.7.3). He thus obtains an infinite system of linear equations

$$b_1+2^{2h+1}b_2+3^{2h+1}b_3+\ldots = (-1)^h f^{(2h+1)}(0) \quad (h = 0, 1, 2,\ldots) \tag{2.8.1}$$

in an infinity of unknowns. It will be observed that all these series are divergent even in the simplest cases: thus $f(x) = x$ has the Fourier series

$$2(\sin x-\tfrac{1}{2}\sin 2x+\tfrac{1}{3}\sin 3x-\ldots), \tag{2.8.2}$$

† Its sum for $x = -\pi$ or π is naturally not $f(x)$ but $\frac{1}{2}\{f(-\pi)+f(\pi)\} = 0$.

and in this case they reduce to

$$1-1+1-\ldots = \tfrac{1}{2}, \qquad 1-2^{2h}+3^{2h}-\ldots = 0 \quad (h = 1, 2, \ldots).$$

We know that these equations are actually true with appropriate definitions, for example, the A definition. But the fact that they are divergent, and that (as we saw in § 2.7) a slight intrusion of divergent series will make 'Fourier's theorem' false, will give an idea of the difficulties in Fourier's way. Judged by modern standards, he was setting himself a hopeless task.

None the less, Fourier's argument is more than an historical curiosity and is still well worth study. Considerable sections of it are correct, or easily restatable so as to become so; it contains ideas important for other purposes; and there are by-products which may still suggest interesting problems.

Let us write (2.8.1) as

$$(2.8.3) \qquad b_1+2^{2h-1}b_2+3^{2h-1}b_3+\ldots = A_h \quad (h = 1, 2, \ldots).$$

Then Fourier's leading idea is to suppress all but the first r equations and all but the first r unknowns, thus obtaining a finite system

$$(2.8.4) \qquad \sum_{n=1}^{r} n^{2h-1}b_n = A_h \quad (h = 1, 2, \ldots, r),$$

to calculate the corresponding values $b_{n,r}$ of the b_n, and to investigate the limit of $b_{n,r}$ when $r \to \infty$. This is now the dominant idea in the theory of the solution of an infinite system of linear equations, and it is in Fourier's work that it appears first.

Fourier, however, does not do exactly this. He varies the A_h as well as the b_n, replacing (2.8.4) by

$$(2.8.5) \qquad \sum_{n=1}^{r} n^{2h-1}b_{n,r} = A_{h,r} \quad (h = 1, 2, \ldots, r).$$

We call this the system (r). Fourier's idea is that we can, by an appropriate choice of the $A_{h,r}$, secure both that $A_{h,r} \to A_h$ and that the $b_{n,r}$ tend to limits b_n.

He tries to show† that if we choose the $A_{h,r}$ so that

$$(2.8.6) \qquad A_{1,1} = A_{1,2} - \frac{A_{2,2}}{2^2},$$

$$(2.8.7) \qquad A_{1,2} = A_{1,3} - \frac{A_{2,3}}{3^2}, \qquad A_{2,2} = A_{2,3} - \frac{A_{3,3}}{3^2},$$

† This part of Fourier's argument is restated in a more accurate form by Darboux in a footnote to p. 191 of the reprint of Fourier's works.

and generally

$$A_{h,r} = A_{h,r+1} - \frac{A_{h+1,r+1}}{(r+1)^2} \quad (h = 1, 2, \dots, r), \tag{2.8.8}$$

then we shall have

$$b_{1,1} = b_{1,2}\left(1-\frac{1}{2^2}\right), \tag{2.8.9}$$

$$b_{1,2} = b_{1,3}\left(1-\frac{1}{3^2}\right), \qquad b_{2,2} = b_{2,3}\left(1-\frac{2^2}{3^2}\right), \tag{2.8.10}$$

and generally

$$b_{n,r} = b_{n,r+1}\left\{1-\frac{n^2}{(r+1)^2}\right\} \quad (n = 1, 2, \dots, r). \tag{2.8.11}$$

It will follow that, if the $A_{h,r}$ satisfy (2.8.8), then $b_{n,r} \to b_n$, where

$$b_n\left\{1-\frac{n^2}{(n+1)^2}\right\}\left\{1-\frac{n^2}{(n+2)^2}\right\}\dots = b_{n,n}, \tag{2.8.12}$$

when $r \to \infty$. We have then to calculate $b_{1,1}$, $b_{2,2}, \dots$ in terms of the A_h (which are *ex hypothesi* the limits of the $A_{h,r}$).

Now $b_{1,1} = A_{1,1}$, by the first of (2.8.5). We express this first in terms of $A_{1,2}$, $A_{2,2}$, by (2.8.6), next in terms of $A_{1,3}$, $A_{2,3}$, $A_{3,3}$, by (2.8.7), and so on. We find that

$$A_{1,1} = A_{1,2} - \frac{A_{2,2}}{2^2} = A_{1,3} - A_{2,3}\left(\frac{1}{2^2}+\frac{1}{3^2}\right) + \frac{A_{3,3}}{2^2 3^2}$$

$$= A_{1,4} - A_{2,4}\left(\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}\right) + A_{3,4}\left(\frac{1}{3^2 4^2}+\frac{1}{4^2 2^2}+\frac{1}{2^2 3^2}\right) - \frac{A_{4,4}}{2^2 3^2 4^2} = \dots,$$

and in the limit

$$b_{1,1} = A_{1,1} = A_1 P_{1,1} - A_2 P_{2,1} + A_3 P_{3,1} - \dots, \tag{2.8.13}$$

where

$$P_{1,1} = 1, \quad P_{2,1} = \sum m_1^{-2}, \quad P_{3,1} = \sum m_1^{-2} m_2^{-2}, \dots, \tag{2.8.14}$$

and the summations are extended over unequal values of m_1, $m_2, \dots$ other than 1. This and (2.8.12) give b_1 in terms of the A_h.

We can calculate $b_{2,2}$, $b_{3,3}, \dots$, and so b_2, $b_3, \dots$, similarly. We express $b_{2,2}$ in terms of $A_{1,2}$, $A_{2,2}$ from the system (2), then in terms of $A_{1,3}$, $A_{2,3}$, $A_{3,3}$ from (2.8.7), and so on. Thus we obtain b_2; and we may obtain b_n similarly by starting from the system (n). The results may be written

$$b_1\left(1-\frac{1}{2^2}\right)\left(1-\frac{1}{3^2}\right)\dots = A_1 P_{1,1} - A_2 P_{2,1} + A_3 P_{3,1} - \dots, \tag{2.8.15}$$

. . . .

$$n b_n \prod_{m \neq n}\left(1-\frac{n^2}{m^2}\right) = A_1 P_{1,n} - A_2 P_{2,n} + A_3 P_{3,n} - \dots, \tag{2.8.16}$$

where

(2.8.17) $\quad P_{1,n} = 1, \qquad P_{2,n} = \sum m_1^{-2}, \qquad P_{3,n} = \sum m_1^{-2} m_2^{-2}, \ldots$

and the summations are now extended over unequal $m_1, m_2, \ldots$ other than n. These equations may be simplified because

$$\prod_{m \neq n} \left(1 - \frac{n^2}{m^2}\right) = \lim_{z \to n} \left\{\frac{\sin \pi z}{\pi z}\left(1 - \frac{z^2}{n^2}\right)^{-1}\right\} = (-1)^{n-1}\tfrac{1}{2}.$$

Thus (2.8.16) becomes

(2.8.18) $\quad (-1)^{n-1}\tfrac{1}{2}nb_n = A_1 P_{1,n} - A_2 P_{2,n} + A_3 P_{3,n} - \ldots.$

It is easy to find $P_{h,n}$ for all h and n. If we write

$$\prod_1^\infty \left(1 - \frac{z^2}{m^2}\right) = \frac{\sin \pi z}{\pi z} = P_1 - P_2 z^2 + P_3 z^4 - \ldots,$$

then $P_{h+1} = \pi^{2h}/(2h+1)!$, and the identity

$$\left(1 - \frac{z^2}{n^2}\right)(P_{1,n} - P_{2,n} z^2 + P_{3,n} z^4 - \ldots) = P_1 - P_2 z^2 + P_3 z^4 - \ldots$$

gives the $P_{h,n}$ in terms of the P_h. Finally, making these calculations and substituting in (2.8.18), we obtain

(2.8.19) $\quad (-1)^{n-1}\tfrac{1}{2}nb_n = A_1 + \left(\frac{1}{n^2} - \frac{\pi^2}{3!}\right)A_2 + \left(\frac{1}{n^4} - \frac{1}{n^2}\frac{\pi^2}{3!} + \frac{\pi^4}{5!}\right)A_3 + \ldots.$

This completes the first and most complex stage of Fourier's argument. Thus if $f(x) = x$, $A_1 = 1$, $A_2 = A_3 = \ldots = 0$, we obtain $\tfrac{1}{2}nb_n = (-1)^{n-1}$ and

$$x = 2(\sin x - \tfrac{1}{2}\sin 2x + \tfrac{1}{3}\sin 3x - \ldots).$$

2.9. Other forms of the coefficients and the series. If we remember that $A_h = (-1)^{h-1} f^{(2h-1)}(0)$, so that

$$f(\pi) = \pi A_1 - \frac{\pi^3}{3!}A_2 + \ldots, \quad f''(\pi) = -\pi A_2 + \frac{\pi^3}{3!}A_3 - \ldots, \quad \ldots,$$

and rearrange (2.8.19) in powers of n^{-2}, we find that

(2.9.1) $\quad b_n = (-1)^{n-1}\dfrac{2}{\pi n}\left\{f(\pi) - \dfrac{f''(\pi)}{n^2} + \dfrac{f''''(\pi)}{n^4} - \ldots\right\}.$

Substituting this series for b_n in $\sum b_n \sin nx$, and rearranging the resulting double series by associating together the terms in $f(\pi), f''(\pi), \ldots$, we obtain

(2.9.2) $\quad \tfrac{1}{2}\pi f(x) = \sum_{h=0}^{\infty} (-1)^h f^{(2h)}(\pi)\left(\sin x - \dfrac{\sin 2x}{2^{2h+1}} + \dfrac{\sin 3x}{3^{2h+1}} - \ldots\right).$

Each of these formulae is interesting in itself, and valid under fairly wide conditions.

We shall have more to say about (2.9.2) in Ch. XIII. Here we are concerned with Fourier's further transformations of (2.9.1). He observes that
$$\chi(x) = f(x) - n^{-2}f''(x) + n^{-4}f''''(x) - \dots$$
satisfies the differential equation
$$n^{-2}\chi''(x) + \chi(x) = f(x),$$
whose general solution is
$$\chi(x) = C\cos nx + D\sin nx + n\sin nx\int\limits_0^x f(t)\cos nt\,dt - n\cos nx\int\limits_0^x f(t)\sin nt\,dt.$$
Since f is odd, χ is odd and $C = \chi(0) = 0$. Hence, putting $x = \pi$, and using (2.9.1), we obtain
$$b_n = (-1)^{n-1}\frac{2}{n\pi}\chi(\pi) = \frac{2}{\pi}\int\limits_0^\pi f(t)\sin nt\,dt,$$
which is the second formula (2.7.2). Thus at last Fourier has arrived at the ordinary formula for the coefficients.

2.10. The validity of Fourier's formulae. It would no doubt be possible to determine conditions on $f(x)$ sufficient to justify all Fourier's elaborate transformations, but a very careful analysis of his argument would be required. Here we shall consider two questions only: (*a*) whether (2.9.1) is in fact a correct formula for b_n, and (*b*) whether the b_n actually satisfy (2.8.1). The second question naturally presupposes some definition of the sums of the divergent series involved.

(1) First, if $f(x)$ is odd and regular along the stretch $-\pi \leqslant x \leqslant \pi$ of the real axis,† we have, since $f(0) = f''(0) = \dots = 0$,
$$\tfrac{1}{2}\pi b_n = \int\limits_0^\pi f(x)\sin nx\,dx = (-1)^{n-1}\left\{\frac{f(\pi)}{n} - \frac{f''(\pi)}{n^3} + \dots + \right.$$
$$\left. + \frac{(-1)^h f^{(2h)}(\pi)}{n^{2h+1}} + \frac{(-1)^h}{n^{2h+1}}\int\limits_0^\pi f^{(2h+1)}(x)\cos nx\,dx\right\},$$
by repeated partial integration. Since the last term in the bracket is $O(n^{-2h-2})$ for large n, we see that the series (2.9.1) is an asymptotic series for b_n.

Next, if
$$|f^{(2h)}(\pi)| < CK^{2h} \quad (h = 0, 1, \dots)$$
for some C and K, then the series is convergent for $n > K$. In particular this will be true if $f(x)$ is an integral function of order 1 and finite type,

† This is, of course, a more general hypothesis than Fourier's: he assumes that the Taylor's series of $f(x)$ is convergent for $|x| \leqslant \pi$.

i.e. if $|f(x)| < De^{L|x|}$ for some D and L. If $L < 1$, in which case also $K < 1$, then the series is convergent for $n \geqslant 1$. In these circumstances (2.9.1) is true in the ordinary sense.

We can also prove that the series is summable, under wider conditions, in various senses, but this demands some knowledge of the definitions of the sum of a divergent series associated with the name of Borel.

(2) We shall now prove that the equations (2.8.1) are correct if the divergent series which they contain are summed by the A method of § 1.3 (2). We again suppose only that $f(x)$ is odd and regular on the stretch $-\pi \leqslant x \leqslant \pi$ of the real axis.

Since $f(x)$ is regular for $-\pi \leqslant x \leqslant \pi$, we have

$$b_n = \frac{1}{\pi i}\int_{-\pi}^{\pi} f(x)e^{nix}\,dx = \frac{1}{\pi i}\int_{C_1} f(x)e^{nix}\,dx,$$

where C_1 is a curve from $-\pi$ to π a little above the real axis. Hence

$$\sum b_n e^{-\delta n} = \frac{1}{\pi i}\int_{C_1} f(x)\sum e^{nix-\delta n}\,dx = \frac{1}{\pi i}\int_{C_1} f(x)\frac{e^{ix-\delta}}{1-e^{ix-\delta}}\,dx,$$

for any positive δ. Differentiating $2h+1$ times with respect to δ, and then replacing the derivative under the integral sign by the corresponding derivative with respect to x, we obtain

$$\sum n^{2h+1}b_n e^{-\delta n} = \frac{(-1)^{h-1}}{\pi}\int_{C_1} f(x)\left(\frac{d}{dx}\right)^{2h+1}\frac{e^{ix-\delta}}{1-e^{ix-\delta}}\,dx.$$

When $\delta \to 0$, the right-hand side tends to

$$\frac{(-1)^{h-1}}{\pi}\int_{C_1} f(x)\left(\frac{d}{dx}\right)^{2h+1}\frac{e^{ix}}{1-e^{ix}}\,dx = \frac{(-1)^{h}}{2\pi i}\int_{C_1} f(x)\left(\frac{d}{dx}\right)^{2h+1}\cot\tfrac{1}{2}x\,dx.$$

Since $f(x)$ is odd, this is half the same integral round C, a complete circuit round the origin in the negative direction; and this, by partial integration, is

$$\frac{(-1)^{h-1}}{2\pi i}\int_{C} f^{(2h+1)}(x)\tfrac{1}{2}\cot\tfrac{1}{2}x\,dx = (-1)^h f^{(2h+1)}(0),$$

which is accordingly the A sum of $\sum n^{2h+1}b_n$. Actually the series are summable by 'Cesàro' methods, $\sum n^{2h+1}b_n$ being summable $(\mathrm{C}, 2h+1)$;† but the A method is the simplest which will sum all of them.

† See § 5.4.

D. *Heaviside's exponential series*

2.11. Heaviside on divergent series. Our last example is one of a different kind, since it comes from quite modern times and from the work of a man who was not a professional mathematician.

Heaviside, in the second volume of his *Electromagnetic theory* (London, 1899), has a long chapter on divergent series. He is plainly not aware that, at the time when this volume was published, a scientific theory of divergent series already existed;† and his work is always unsystematic and often obscure. He does not attempt to develop anything which can be called a 'theory' of divergent series, his attitude towards them being, at bottom, that of Euler 150 years before: indeed Euler had the clearer ideas. But Heaviside, whatever his merits as a mathematician, was a man of much talent and originality, and what he says (if often irritating to a mathematician) is always interesting.

It may be advisable to substantiate these assertions by quotations from Heaviside's writings.

'I must say a few words on the subject of generalized differentiation and divergent series. . . . It is not easy to get up any enthusiasm after it has been artificially cooled by the wet blankets of rigorists. . . . I have been informed that I have been the means of stimulating some interest in the subject. Perhaps not in England to any extent worth speaking of, but certainly in Paris it is a fact that a big prize has been offered lately on the subject of the part played by divergent series in analysis. . . . I hope the prize-winner will have something substantial to say. . . .

'In *O.P.M.*‡ I have stated the growth of my views about divergent series up to that time. . . . I have avoided defining the meaning of equivalence. The definitions will make themselves in time. . . . My first notion of a series was that to have a finite value it must be convergent. . . . A divergent series also, of course, has an infinite value. Solutions of physical problems must always be in finite terms or convergent series, otherwise nonsense is made. . . .

'Then came a partial removal of ignorant blindness. In some physical problems divergent series are actually used, notably by Stokes, referring to the divergent formula for the oscillating function $J_n(x)$. He showed that the error was less than the last term included. Now here the terms are alternately positive and negative. This seems to give a clue. . . .

'There are certainly three kinds of equivalence. . . .§ Equivalence does not mean identity. . . . But the numerical meaning of divergent series still remains obscure.

† Borel's memoirs on divergent series were published during the years 1895–9, his book in 1901. Poincaré's theory of asymptotic series dated from 1886.

‡ 'On Operators in Physical Mathematics': a series of three papers presented to the Royal Society during 1892–4 but never printed in full.

§ Numerical, analytical, algebraical. Heaviside means, of course, that

$$1+x+x^2+\ldots = (1-x)^{-1}$$

may mean (*a*) that $1+x+x^2+\ldots$ converges to $(1-x)^{-1}$, (*b*) that it is a 'representation' of the function $(1-x)^{-1}$, (*c*) that it is the result of the algebraical process of 'long division' of 1 by $1-x$. Euler would have said the same.

. . . There will have to be a theory of divergent series, or say a larger theory of functions than the present, including convergent and divergent series in one harmonious whole. . . .' (*Electromagnetic theory*, 2, 434–50.)

The 'rigorists' whom Heaviside disliked so much had provided what he asked for, even at the time when he wrote.

2.12. The generalized exponential series. There is one particular series which Heaviside uses freely, and which he seems to have been the first to use, though it is a special case of one stated many years before by Riemann. This is the series

$$S = S(x,c) = \sum_{r=-\infty}^{\infty} \frac{x^{c-r}}{\Gamma(c-r+1)}, \tag{2.12.1}$$

where $x > 0$, c is real, and the coefficient is to be taken as 0 if c is an integer n and $r > n$: in this case S reduces to the ordinary exponential series. Otherwise S is divergent for all x; but, since it reproduces itself when differentiated formally, it is natural to suppose that it should have the sum e^x, in some sense, for all c.

We suppose that c is non-integral, R integral, and $R > c$. Then it is easily verified by partial integration, or by differentiation of the result, that

$$\frac{1}{\Gamma(c-R)} \int_x^\infty e^{-t}t^{c-R-1}\,dt = 1-e^{-x}\sum_{n=0}^{\infty} \frac{x^{c-R+n}}{\Gamma(c-R+n+1)}.$$

Hence

$$S_R(x,c) = \sum_{r=-\infty}^{R} \frac{x^{c-r}}{\Gamma(c-r+1)} = \sum_{n=0}^{\infty} \frac{x^{c-R+n}}{\Gamma(c-R+n+1)}$$

$$= e^x - \frac{e^x}{\Gamma(c-R)} \int_x^\infty e^{-t}t^{c-R-1}\,dt = e^x + Q_R,$$

say. The sign of Q_R is that of $-\Gamma(c-R)$, and

$$|Q_R| < \frac{1}{|\Gamma(c-R)|} \int_x^\infty t^{c-R-1}\,dt = \frac{x^{c-R}}{|\Gamma(c-R+1)|}.$$

The signs of the terms in S with $r > c$ alternate in sign. If, for example, u_R, the last term in S_R, is positive, then $\Gamma(c-R) < 0$, $0 < Q_R < u_R$, and $e^x < S_R < e^x+u_R$. If u_R is negative, then

$$e^x+u_R < S_R < e^x.$$

Thus the series S represents e^x asymptotically in a sense analogous to that of § 2.5; its terms, from a certain point on, alternate in sign; and the error involved in stopping at any term is of the same sign as, and numerically less than, the last term retained.

2.13. The series $\sum \phi^{(r)}(x)$. Heaviside's series is a special case of the series

$$S = S(x) = \sum_{-\infty}^{\infty} \phi^{(r)}(x), \tag{2.13.1}$$

where $\phi^{(r)}(x)$, the rth generalized derivative of $\phi(x)$, is defined for $r = -s < 0$ by

$$\phi^{(-s)}(x) = \phi_s(x) = \left(\int_0^x dt\right)^s \phi(t) = \frac{1}{(s-1)!}\int_0^x (x-t)^{s-1}\phi(t)\,dt, \tag{2.13.2}$$

and for $r = -s+N$ as the Nth differential coefficient of $\phi_s(x)$. If $\phi(x)$ is a multiple of x^c, and $c > -1$, then (2.13.1) reduces to (2.12.1).

If

$$S = \Big(\sum_{r<0} + \sum_{r\geqslant 0}\Big)\phi^{(r)}(x) = S^{(1)}+S^{(2)}, \tag{2.13.3}$$

say, then

$$S^{(1)} = \sum_{s>0} \frac{1}{(s-1)!}\int_0^x (x-t)^{s-1}\phi(t)\,dt = \int_0^x e^{x-t}\phi(t)\,dt \tag{2.13.4}$$

for any integrable ϕ.

Let us assume for simplicity that $\phi(t)$ is indefinitely differentiable throughout any finite interval of positive t, and that

$$\int_0^\infty e^{-t}\phi^{(r)}(t)\,dt$$

is convergent for $r = 0, 1,\ldots$, in which case $e^{-t}\phi^{(r)}(t) \to 0$, when $t \to \infty$, for $r = 0, 1,\ldots$. Then

$$\int_x^\infty e^{-t}\phi(t)\,dt - \int_x^\infty e^{-t}\phi^{(R+1)}(t)\,dt = e^{-x}\sum_{r=0}^{R} \phi^{(r)}(x),$$

by partial integration; and so

$$S_R^{(2)} = \sum_{r=0}^{R} \phi^{(r)}(x) = e^x\left\{\int_x^\infty e^{-t}\phi(t)\,dt - \int_x^\infty e^{-t}\phi^{(R+1)}(t)\,dt\right\}.$$

Combining this with (2.13.4), we find

$$S_R = \sum_{r=-\infty}^{R} \phi^{(r)}(x) = e^x\int_0^\infty e^{-t}\phi(t)\,dt - e^x\int_x^\infty e^{-t}\phi^{(R+1)}(t)\,dt. \tag{2.13.5}$$

If now $|\phi^{(R+1)}(x)| < \chi_{R+1}(x)$, and $\chi_{R+1}(x) \to 0$, for every R, when $x \to \infty$, then (2.13.5) gives

$$S_R = e^x\int_0^\infty e^{-t}\phi(t)\,dt + O\{\chi_{R+1}(x)\} = Ae^x + O\{\chi_{R+1}(x)\},$$

and in this sense S is an asymptotic series for Ae^x.

2.14. The generalized binomial series. Heaviside has also a 'generalized binomial series', viz.

$$(2.14.1.)\quad (1+x)^n = \sum_{s=-\infty}^{\infty} x^{n-m+s}\frac{\Gamma(n+1)}{\Gamma(m-s+1)\Gamma(n+s-m+1)},$$

where m and n are not usually integral. This series, unlike the exponential series, appears explicitly in Riemann's earlier work.

If m and n are integral and n positive, then (2.14.1) reduces to the elementary binomial theorem; if m only is integral, to the ordinary infinite series,

$$\sum_{s=-\infty}^{m} x^{n-m+s}\frac{\Gamma(n+1)}{\Gamma(m-s+1)\Gamma(n+s-m+1)} = \sum_{r=0}^{\infty} x^{n-r}\frac{\Gamma(n+1)}{\Gamma(r+1)\Gamma(n-r+1)},$$

which converges to $x^n(1+x^{-1})^n = (1+x)^n$ if $|x| > 1$. Generally the series is infinite at both ends, and convergent at one end, divergent at the other, according to the value of x.

If we separate the positive and negative values of s, and write the two resulting series at length in the notation usual for hypergeometric series, we obtain

$$(2.14.2)\qquad \frac{\Gamma(m+1)\Gamma(n-m+1)}{\Gamma(n+1)}x^{m-n}(1+x)^n$$

$$= F(1,-m,n-m+1,-x)+\frac{n-m}{(m+1)x}F\left(1,-n+m+1,m+2,-\frac{1}{x}\right).$$

If, for example, $0 < x < 1$, then the first series on the right is convergent; the second is divergent, but summable in various ways, and represents the analytic continuation of the function which it defines when convergent. The formula may be proved directly or deduced from known theorems concerning the relations between different hypergeometric functions.

NOTES ON CHAPTER II

§ 2.2. Euler, 'Remarques sur un beau rapport entre les séries des puissances tant directes que réciproques', *Histoire de l'Académie des Sciences et Belles-lettres* (Mémoires de l'Académie), 17 (Berlin, 1768), 83–106 [*Opera* (I), 15, 70–91]. The volume covers the year 1761, and the paper had been read in 1749.

Cahen, *AEN* (3), 11 (1894), 75–164 (75–6), seems to have been the first modern writer to call attention to Euler's paper. Landau, *Bibliotheca Math.* (3), 7 (1906), 69–79, gives a full account of it, with the appropriate references to other writers. It seems that no one before Riemann (1859) gave a satisfactory proof of (2.2.2), but that Schlömilch had stated (2.2.7) in 1849 and proved it in 1858. The standard proofs of (2.2.2) are given in Landau, *Handbuch*, 281–98: see also Ingham, 41–8,

Whittaker and Watson, 268–9. Many other proofs have been given by other writers.

§ 2.3. For (2.3.2) see, e.g., Bromwich, 298.

§ 2.4. Euler's discussions of the series (2.4.1) seem to have begun in his correspondence with N. Bernoulli: see in particular *Opera* (I), 14, 585. Other references will be found in Reiff's book quoted in the note on § 1.3. The summability of the series by various methods is discussed by Hardy, *PCPS*, 37 (1941), 1–8: see § 8.11.

There is a systematic account of the theory of $\operatorname{li} e^{-x}$ in Nielsen, *Theorie des Integral-logarithmus und verwandter Transzendenten* (Leipzig, 1906).

§ 2.5. Poincaré's memoir was published in *AM*, 8 (1886), 295–344. There are accounts of the theory of asymptotic series in Borel, ch. 1; Bromwich, ch. 12; Knopp, ch. 14; and Ford, *Studies*.

For the theorems of Watson and Carleman see Watson, *PTRS*(A), 211 (1912), 279–313; Carleman, *Les fonctions quasi-analytiques* (Paris, 1926); and § 8.11.

§ 2.6. Lacroix, *Traité du calcul*, 3, ed. 2 (Paris, 1819), 346–8; Bromwich, 336.

§ 2.7. There are short accounts of the relevant parts of the theory of Fourier series in Hardy and Rogosinski and in other books there referred to, and a very full one in Zygmund.

The fullest account of the early history of the formulae (2.7.2) is that in Burkhardt's *Enzyklopädie* article quoted under § 1.3.

§ 2.8. Fourier, *Théorie analytique de la chaleur*, ed. 2 (Paris, 1822), 187 et seq. (reprinted in vol. 1 of his *Œuvres*). There are short accounts of Fourier's analysis in F. Riesz, *Les systèmes d'équations linéaires à une infinité d'inconnues* (Paris, 1913), ch. 1, and Hardy, *Annals*, 36 (1935), 167–81; but both are condensed, and neither author quite does justice to Fourier.

Dr. Bosanquet observes (1) that it is at any rate doubtful whether it is always possible, under Fourier's conditions, to choose $A_{h,r}$ so as to satisfy (2.8.8) and $A_{h,r} \to A_h$; (2) that we can deduce directly from (2.8.5) that

$$nb_{n,r}\prod_{m=1}^{r}{}'\left(1-\frac{n^2}{m^2}\right)=\left\{\prod_{m=1}^{r}{}'\left(1-\frac{E}{m^2}\right)\right\}A_{1,r}=\sum_{h=1}^{r}(-1)^{h-1}A_{h,r}P_{h,n}^{(r)},$$

where the dash implies the omission of the value $m = n$, $EA_{h,r} = A_{h+1,r}$, and

$$P_{1,n}^{(r)}=1,\quad P_{h,n}^{(r)}=\sum m_1^{-2}m_2^{-2}\dots m_{h-1}^{-2}\quad (h>1),$$

the summation extending over unequal $m_1, m_2, \dots$ from 1 to r other than n. If we then make $r \to \infty$, and suppose that $A_{h,r} \to A_h$, we obtain (2.8.16) without using the special relations (2.8.6)–(2.8.12).

§ 2.10(2). The argument here may be generalized: see Hardy, l.c. under § 2.8, 172. Very general theorems concerning the Cesàro summability of derived series of Fourier series, of which the series considered here are special examples, were proved by W. H. Young, *PLMS* (2), 17 (1918), 195–236. Still more general results, and full references, will be found in Zygmund, 257 *et seq*, and in Bosanquet, *PLMS* (2), 46 (1940), 270–89.

§ 2.12. Heaviside's exponential series, and the generalized binomial series of § 2.14, are both special cases of a generalized form of Taylor's series which occurs on p. 335 of Riemann's posthumous fragment 'Versuch einer allgemeinen Auffassung der Integration und Differentiation' (*Werke*, 331–44). Riemann's expansion is

$$f(x+h)=\sum\frac{h^{m+r}}{\Gamma(m+r+1)}D^{m+r}f(x),$$

where r is fixed, and in general non-integral, m runs from $-\infty$ to ∞, and D^{m+r} is a symbol of generalized differentiation. Riemann does not write down the exponential series explicitly, and makes no attempt at a rigorous discussion. For this see Hardy, *JLMS*, 20 (1945), 48–57.

The fragment is taken from a manuscript dated 14 Jan. 1847, when Riemann was a student. As the editors (Dedekind and Weber) remark, it was never intended for publication; but it contains the first definition of 'Riemann-Liouville' integrals, and no doubt marks the beginning of Riemann's work on hypergeometric series.

The asymptotic character of the series (2.12.1) was proved by Barnes, *TCPS*, 20 (1908), 253–79. Barnes's proof is valid for complex x with $|\arg x| < \pi$.

§ 2.13. Pólya has proved that if the series (2.13.1) is convergent for any x for which $\phi(x)$ is regular, then $\phi(x)$ is an integral function, and the series is uniformly convergent in any bounded region of x (so that its sum is necessarily a multiple of e^x). See Pólya and Szegö, 1, 133 and 314.

Hardy, l.c. under § 2.8, discusses the summability of (2.13.1) by methods of Borel's type.

§ 2.14. The binomial series occurs, as formula (3), on p. 336 of Riemann's fragment. He comments on its failure for negative integral n.

The formula (2.14.2) occurs, for example, in Barnes, *PLMS* (2), 6 (1908), 141–77 (146, formula I). It may be proved directly by integrating

$$\int (-u)^{-m-1}(1-u)^n \frac{du}{1+xu}$$

round an appropriate contour.

III

GENERAL THEOREMS

3.1. Generalities concerning linear transformations. The theory of divergent series is concerned with generalizations of the notion of the limit of a sequence (s_n), which are usually effected by an auxiliary sequence of linear means of s_n. Thus in § 1.3 we defined the (C, 1) limit of (s_n), or the (C, 1) sum of $\sum a_n$, as the limit of

(3.1.1) $$t_m = \frac{s_0+s_1+...+s_m}{m+1}$$

when $m \to \infty$; and the A limit of (s_n), or the A sum of $\sum a_n$, as the limit of

(3.1.2) $$t(x) = \sum a_n x^n = \sum x^n(1-x)s_n$$

when $x \to 1$ through values less than 1. In each case the auxiliary means are of the form

(3.1.3) $$t_m = \sum c_{m,n} s_n \quad (m = 0, 1, 2,...)$$

or

(3.1.4) $$t(x) = \sum c_n(x)s_n,$$

where x is a continuous parameter.† Thus in (3.1.1)

$$c_{m,n} = (m+1)^{-1} \quad (0 \leqslant n \leqslant m), \qquad c_{m,n} = 0 \quad (n > m),$$

and in (3.1.2) $c_n(x) = x^n(1-x)$. In one case they depend on an integral parameter m, in the other on a continuous parameter x; but, as we shall see, this distinction is not very important. For the moment we consider means of the type (3.1.3).

We call the system of equations (3.1.3), which we may write shortly as

(3.1.5) $$t = \mathrm{T}(s),$$

a *linear transformation* T; t_m the *transform* of s_n by T; and the matrix

$$|\mathrm{T}| = (c_{m,n}),$$

in which $c_{m,n}$ is the element in the mth row and nth column, the *matrix of* T.

† Summations are over 0, 1, 2,... when there is no indication to the contrary. The variable of summation will not be shown explicitly unless this is necessary to avoid ambiguity: it is obvious, for example, that in (3.1.3) and (3.1.4) the summation must be with respect to n, so that, for example, $\Sigma\, c_{m,n}\, s_n$ means

$$\sum_{n=0}^{\infty} c_{m,n}\, s_n.$$

3.2. Regular transformations. The most important transformations are *regular*. We say that T is regular if

$$t_m \to s \quad (m \to \infty) \tag{3.2.1}$$

whenever

$$s_n \to s \quad (n \to \infty). \tag{3.2.2}$$

We regard the first assertion as including that of the existence of t_m for every m, i.e. the convergence of all the series (3.1.3). Thus, after Cauchy's theorem quoted in § 1.4, the transformation (3.1.1) is regular.

There is an important theorem, due to Toeplitz and Schur, which states necessary and sufficient conditions for the regularity of T. We prove this theorem (Theorem 2) in § 3.3: it is convenient to associate it with two other theorems of a similar character concerning different classes of transformations. We call the class of linear transformations $\mathfrak{T}$, the class of regular transformations $\mathfrak{T}_r$. The class $\mathfrak{T}_c$ is the class of transformations which transform all convergent sequences into convergent sequences, i.e. transformations such that the convergence of s_n to s implies the convergence of t_m to *some* limit t. Thus $\mathfrak{T}_r$ is the subclass of $\mathfrak{T}_c$ in which t is necessarily the same as s. The class $\mathfrak{T}_c^*$ is the class of transformations which convert all bounded sequences into convergent sequences, i.e. transformations such that $s_n = O(1)$ implies $t_m \to t$. It is plain that $\mathfrak{T}_c^*$ is also a subclass of $\mathfrak{T}_c$; but $\mathfrak{T}_c^*$ and $\mathfrak{T}_r$ are, as we shall see, mutually exclusive. We shall prove the following three theorems.

THEOREM 1. *In order that* T *should belong to* $\mathfrak{T}_c$, *it is necessary and sufficient* (i) *that*

$$\gamma_m = \sum |c_{m,n}| < H, \tag{3.2.3}$$

where H is independent of m; (ii) *that*

$$c_{m,n} \to \delta_n, \tag{3.2.4}$$

for each n, when $m \to \infty$; (iii) *that*

$$c_m = \sum c_{m,n} \to \delta \tag{3.2.5}$$

when $m \to \infty$. In these circumstances $\sum \delta_n$ is absolutely convergent, and

$$t_m \to t = \delta s + \sum \delta_n(s_n - s) = s(\delta - \sum \delta_n) + \sum \delta_n s_n, \tag{3.2.6}$$

when $m \to \infty$, whenever $s_n \to s$.

Here, of course, the limits δ_n and δ are finite.

THEOREM 2. *In order that* T *should belong to* $\mathfrak{T}_r$ (*i.e. that* T *should be regular*), *it is necessary and sufficient that the conditions of Theorem* 1 *should be satisfied, that $\delta_n = 0$ for each n, and that $\delta = 1$.*

THEOREM 3. *In order that* T *should belong to* $\mathfrak{T}_c^*$, *it is necessary and sufficient that* $c_{m,n} \to \delta_n$ *for each* n, *i.e. that the second condition of Theorem* 1 *should be satisfied, and that the series* $\sum |c_{m,n}|$ *should converge uniformly in* m. *In these circumstances the first and third conditions of Theorem* 1 *are necessarily satisfied,* $\sum \delta_n = \delta$, *and*

$$t_m \to t = \sum \delta_n s_n$$

for all bounded sequences (s_n).

3.3. Proof of Theorems 1 and 2. (1) We prove first that the conditions of Theorem 1 are sufficient. Since $s_n \to s$, s_n is bounded, and it follows from (3.2.3) that all the series (3.1.3) are absolutely convergent.

Next, the series (3.2.5) are absolutely convergent. Also, by (3.2.4),

$$\sum_0^N |\delta_n| = \lim_{m\to\infty} \sum_0^N |c_{m,n}| \leqslant H$$

for every N, so that

(3.3.1) $$\sum |\delta_n| \leqslant H.$$

Thus $\sum \delta_n$, $\sum \delta_n s_n$, and the other series in (3.2.6) are absolutely convergent.

Suppose first that $s = 0$. Then we can choose $N = N(\epsilon)$ so that

(3.3.2) $$|s_n| < \epsilon/4H \quad (n > N).$$

Now

$$t_m - \sum \delta_n s_n = \sum (c_{m,n}-\delta_n)s_n = \sum_0^N (c_{m,n}-\delta_n)s_n + \sum_{N+1}^\infty (c_{m,n}-\delta_n)s_n = U+V,$$

say. Here $$|V| \leqslant \frac{\epsilon}{4H} \sum_{N+1}^\infty (|c_{m,n}|+|\delta_n|) \leqslant \tfrac{1}{2}\epsilon,$$

by (3.2.3), (3.3.1), and (3.3.2); and $U \to 0$ when N is fixed and $m \to \infty$ by (3.2.4), so that $|U| < \frac{1}{2}\epsilon$ for $m \geqslant M(\epsilon, N) = M(\epsilon)$. Hence

$$|t_m - \sum \delta_n s_n| < \epsilon$$

for $m \geqslant M(\epsilon)$, and $t_m \to \sum \delta_n s_n$. Thus the conditions (3.2.3) and (3.2.4), without (3.2.5), are sufficient when $s = 0$.

In the general case we write

$$s'_n = s_n - s, \qquad t'_m = \sum c_{m,n} s'_n.$$

Then $s'_n \to 0$ and therefore

$$t'_m \to \sum \delta_n s'_n.$$

Hence, now using (3.2.5),

$$t_m = \sum c_{m,n}(s'_n+s) = t'_m + sc_m \to \sum \delta_n s'_n + \delta s = s(\delta - \sum \delta_n) + \sum \delta_n s_n = t.$$

Thus the conditions of the theorem are sufficient in any case.

(2) We have now to prove the conditions necessary. We suppose that T belongs to $\mathfrak{T}_c$.

(a) Take $s_k = 1$, $s_n = 0$ when $n \neq k$, so that $s_n \to 0$. Then $t_m = c_{m,k}$, and therefore $c_{m,k}$ tends to a limit δ_k when $m \to \infty$. Thus (3.2.4) is a necessary condition.

(b) Take $s_n = 1$ for all n, so that $s_n \to 1$. Then

$$t_m = \sum c_{m,n} = c_m,$$

and therefore c_m tends to a limit δ. Thus (3.2.5) is necessary.

(c) It remains to prove (3.2.3) necessary. This is the main point of the theorem.

First, γ_m is finite for every m. For if $\gamma_m = \infty$ we can choose (ϵ_n) so that

$$\epsilon_n > 0, \qquad \epsilon_n \to 0, \qquad \sum \epsilon_n |c_{m,n}| = \infty.\dagger$$

If then we take $s_n = \epsilon_n \operatorname{sgn} \overline{c_{m,n}}$, we have $s_n \to 0$ and

$$t_m = \sum \epsilon_n |c_{m,n}| = \infty,$$

in contradiction to our hypotheses.

Thus γ_m is finite for every m, and we have to prove γ_m bounded. If not then, given G, we can find an m such that $\gamma_m > G$. We write

(3.3.3) $$\gamma_{m,n} = \sum_{\nu=0}^{n} |c_{m,\nu}|,$$ (3.3.4) $$d_n = \sum_{\nu=0}^{n} |\delta_\nu|.$$

We know already that $\gamma_{m,n} \to \gamma_m$ when $n \to \infty$, and that $c_{m,\nu} \to \delta_\nu$ (so that $\gamma_{m,n} \to d_n$) when $m \to \infty$.

Starting from an arbitrary n_1, we construct two increasing sequences $m_1, m_2, m_3, \ldots$ and $n_1, n_2, n_3, \ldots$. We suppose that $m_1, m_2, \ldots, m_{r-1}, n_1, n_2, \ldots, n_r$ have been determined already, and choose m_r and n_{r+1} as follows. Since $\gamma_m > G$ for any G and some m, we can choose $m_r > m_{r-1}$ so that

(3.3.5) $$\gamma_{m_r} = \sum |c_{m_r,n}| > 2rd_{n_r} + r^2 + 2r + 2.$$

Since $$\sum_{0}^{n_r} |c_{m,n}| \to \sum_{0}^{n_r} |\delta_n| = d_{n_r}$$

when $m \to \infty$, we can suppose also that

(3.3.6) $$\gamma_{m_r,n_r} = \sum_{0}^{n_r} |c_{m_r,n}| < d_{n_r} + 1.$$

† For example, we may, with Abel, take $\epsilon_n = \left(\sum_{\nu=N}^{n} |c_{m,\nu}|\right)^{-\frac{1}{2}}$, where $c_{m,N}$ is the first $c_{m,\nu} \neq 0$.

Since $\gamma_{m_r,n} \to \gamma_{m_r}$ when $n \to \infty$, we can then choose $n_{r+1} > n_r$ so that

(3.3.7) $$\gamma_{m_r} - \gamma_{m_r,n_{r+1}} = \sum_{n_{r+1}+1}^{\infty} |c_{m_r,n}| < 1.$$

It then follows from (3.3.5)–(3.3.7) that

(3.3.8) $$\sum_{n=n_r+1}^{n_{r+1}} |c_{m_r,n}| > rd_{n_r} + r^2 + 2r.$$

We now take

$$s_n = 0 \quad (n \leqslant n_1), \qquad s_n = r^{-1}\operatorname{sgn}\overline{c_{m_r,n}} \quad (n_r < n \leqslant n_{r+1})$$

for $r = 1, 2,\ldots$. Then $|s_n| \leqslant 1$, $s_n \to 0$, and

$$|t_{m_r}| \geqslant \frac{1}{r}\sum_{n_r+1}^{n_{r+1}} |c_{m_r,n}| - \sum_{0}^{n_r} |c_{m_r,n}| - \sum_{n_{r+1}+1}^{\infty} |c_{m_r,n}|$$
$$> r^{-1}(rd_{n_r} + r^2 + 2r) - (d_{n_r} + 1) - 1 = r.$$

Hence $t_m \to \infty$ when $m \to \infty$ through the sequence (m_r), and T is not a transformation of $\mathfrak{T}_c$. The contradiction completes the proof of Theorem 1.

The only point of the proof which presents difficulty is that of the necessity of the condition (3.2.3). This may be elucidated as follows. Suppose that we wished to prove (3.2.3) a necessary condition for the truth of the implication

$$|s_n| \leqslant K \rightarrow |t_m| \leqslant HK,$$

a theorem about *uniform boundedness* instead of about *convergence*. We should take a fixed m and define s_n by $s_n = K\operatorname{sgn}\overline{c_{m,n}}$. Then

$$t_m = K\sum |c_{m,n}| = K\gamma_m,$$

and (3.2.3) follows. The proof of Theorem 1 depends (at the critical point) on a combination of this device with the use of 'rapidly increasing' sequences. Such proofs are common, for example, in the theory of the 'convergence defects' of Fourier series.

It is now easy to prove Theorem 2. First, the conditions are sufficient because they include those of Theorem 1, so that

$$t_m \to t = \delta s + \sum \delta_n(s_n - s) = s.$$

Secondly, the proof of their necessity is the same: the proof of (3.2.3), indeed, is a little simplified because $\delta_n = 0$ and so $d_n = 0$.

3.4. Proof of Theorem 3. If the conditions of Theorem 3 are satisfied, and s_n is bounded, then $\sum c_{m,n}$, $\sum c_{m,n} s_n$ are uniformly convergent. Hence

$$\lim t_m = \lim \sum c_{m,n} s_n = \sum (\lim c_{m,n}) s_n = \sum \delta_n s_n$$

when $m \to \infty$, and the conditions are sufficient. In particular, taking $s_n = 1$,

$$\delta = \lim \sum c_{m,n} = \sum \delta_n.$$

The conditions (3.2.4) and (3.2.5), being necessary in Theorem 1, are *a fortiori* necessary here. It remains to prove that

(3.4.1) $$\gamma_m = \sum |c_{m,n}|$$

(which is certainly bounded) must be uniformly convergent.

We show first that it is sufficient to prove this in the special case in which $\delta_n = 0$ for every n. If T belongs to $\mathfrak{T}_c^*$ it belongs to $\mathfrak{T}_c$, so that $\sum |\delta_n| < \infty$. The equations

$$t'_m = \sum (c_{m,n} - \delta_n)s_n = \sum c'_{m,n} s_n$$

define a transformation T′ for which $c'_{m,n} \to 0$ when $m \to \infty$. If T belongs to $\mathfrak{T}_c^*$ and $s_n = O(1)$, then $t_m \to t$ and

$$t'_m \to t - \sum \delta_n s_n,$$

so that T′ also belongs to $\mathfrak{T}_c^*$. Hence, if the conclusion has been established in the special case, $\sum |c'_{m,n}|$ is uniformly convergent, and therefore $\sum |c_{m,n}| = \sum |c'_{m,n} + \delta_n|$ is uniformly convergent.

We observe next that the condition of uniform convergence may be stated in a different form by use of (3.2.4). If (3.4.1) is uniformly convergent then

(3.4.2) $$\gamma_m = \sum |c_{m,n}| \to \sum |\delta_n|;$$

and the converse is also true, by a well-known theorem of Dini, because $|c_{m,n}| \geq 0$.† Thus we may replace the condition of uniform convergence by (3.4.2); and, in the special case which it is sufficient to consider, this condition reduces to

(3.4.3) $$\gamma_m = \sum |c_{m,n}| \to 0.$$

† The substance of the theorem, at any rate, is Dini's, but he stated it in a rather different form (for uniform convergence over an interval of values of a continuous variable). It may therefore be advisable to insert an explicit proof of what is actually wanted here, viz. that *if* $u_{m,n} \geq 0$, $u_{m,n} \to U_n$ *when* $m \to \infty$, $\Sigma u_{m,n}$ *and* ΣU_n *are convergent, and*

$$\Sigma u_{m,n} \to \Sigma U_n$$

when $m \to \infty$, *then* $\Sigma u_{m,n}$ *converges uniformly in* m.

In fact

$$\sum_{N+1}^{\infty} u_{m,n} = \left(\sum_0^{\infty} u_{m,n} - \sum_0^{\infty} U_n\right) + \sum_{N+1}^{\infty} U_n - \sum_0^{N} (u_{m,n} - U_n) = P + Q + R,$$

say, so that
$$0 \leq \sum_{N+1}^{\infty} u_{m,n} \leq |P| + |Q| + |R|.$$

We can choose $N(\epsilon)$ so that $|Q| < \epsilon$; and, when $N(\epsilon)$ is fixed, we can choose $M(\epsilon, N) = M(\epsilon)$ so that $|P| < \epsilon$ and $|R| < \epsilon$ for $m \geq M(\epsilon)$. Thus

(a) $$0 \leq \sum_{N+1}^{\infty} u_{m,n} < 3\epsilon$$

for $m \geq M(\epsilon)$ and $N = N(\epsilon)$, and therefore (since $u_{m,n} \geq 0$) for $m \geq M(\epsilon)$, $N \geq N(\epsilon)$. But when $M(\epsilon)$ is fixed we can choose $N_1(\epsilon) \geq N(\epsilon)$ so that (a) is also true for $0 \leq m < M(\epsilon)$ and $N \geq N_1(\epsilon)$, and therefore true for $N \geq N_1(\epsilon)$ and all m.

We have therefore to prove that (3.4.3) is a necessary condition for a transformation T, with $\delta_n = 0$, to belong to $\mathfrak{T}_c^*$.

If (3.4.3) is false, there is a number $\gamma > 0$, and a sequence $(m^{(i)})$, such that

$$\gamma_m = \sum |c_{m,n}| \to \gamma \tag{3.4.4}$$

when $m = m^{(i)}$ and $i \to \infty$. We shall then define a bounded sequence (s_n) such that t_m does not tend to a limit when $m \to \infty$ through $(m^{(i)})$.

We construct increasing sequences (m_r) and (n_r), the first a subsequence of $(m^{(i)})$, as follows. Suppose that $m_1, m_2,..., m_{r-1}$, and $n_1, n_2,..., n_r$ have been determined. Since $\gamma_m \to \gamma$ and $c_{m,n} \to 0$ when $m \to \infty$, we can choose an $m_r > m_{r-1}$, in $(m^{(i)})$, so that

$$|\gamma_{m_r} - \gamma| < 2^{-r}, \tag{3.4.5}$$

$$\sum_0^{n_r} |c_{m_r,n}| < 2^{-r}. \tag{3.4.6}$$

Since $\sum |c_{m_r,n}|$ is convergent, we can then choose $n_{r+1} > n_r$ so that

$$\sum_{n_{r+1}+1}^{\infty} |c_{m_r,n}| < 2^{-r}; \tag{3.4.7}$$

and it follows from (3.4.4)–(3.4.7) that

$$\left|\sum_{n_r+1}^{n_{r+1}} |c_{m_r,n}| - \gamma\right| < 3.2^{-r}. \tag{3.4.8}$$

We now define s_n by

$$s_n = 0 \quad (n \leqslant n_1), \qquad s_n = (-1)^r \operatorname{sgn} \overline{c_{m_r,n}} \quad (n_r < n \leqslant n_{r+1}) \tag{3.4.9}$$

for $r = 1, 2,...$. Then $|s_n| \leqslant 1$, and

$$\left|\sum_0^{n_r} c_{m_r,n}\, s_n\right| < 2^{-r}, \qquad \left|\sum_{n_{r+1}+1}^{\infty} c_{m_r,n}\, s_n\right| < 2^{-r}, \tag{3.4.10}$$

by (3.4.6) and (3.4.7). Also

$$\sum_{n_r+1}^{n_{r+1}} c_{m_r,n}\, s_n = (-1)^r \sum_{n_r+1}^{n_{r+1}} |c_{m_r,n}|$$

by (3.4.9); and so

$$\left|\sum_{n_r+1}^{n_{r+1}} c_{m_r,n}\, s_n - (-1)^r \gamma\right| < 3.2^{-r} \tag{3.4.11}$$

by (3.4.8). Finally, by (3.4.10) and (3.4.11),

$$|t_{m_r} - (-1)^r \gamma| < 5.2^{-r};$$

and therefore t_{m_r} does not tend to a limit. This completes the proof of Theorem 3.

3.5. Variants and analogues. There are many variants of the theorems of § 3.2, which we shall not attempt to enumerate systematically. We mention only a few which will be useful to us later.

(1) The first concerns sequences which tend to zero.

THEOREM 4. *In order that $s_n \to 0$ should imply $t_m \to 0$, it is necessary and sufficient that condition* (3.2.3) *of Theorem* 1 *should be satisfied and that $c_{m,n}$ should tend to* 0 *for each n.*

The sufficiency of the conditions follows from the argument of § 3.3 (1), with $\delta_n = 0$ and $s = 0$. In this case the condition that c_m should tend to a limit is not wanted. The argument of § 3.3 (2), (*a*) and (*c*), also shows that the two conditions retained are necessary, but that of (*b*) is inapplicable.

(2) There are analogues in which m is replaced by a continuous parameter x. Thus the analogue of Theorem 2 is

THEOREM 5. *Suppose that x is a continuous parameter which tends to infinity, and that*

(3.5.1) $$t(x) = \sum c_n(x) s_n.$$

Then the conditions (i) *that $\sum |c_n(x)|$ should be convergent for $x \geqslant 0$, and*

(3.5.2) $$\sum |c_n(x)| < H,$$

where H is independent of x, for $x \geqslant x_0$; (ii) *that*

(3.5.3) $$c_n(x) \to 0,$$

when $x \to \infty$, for every n; and (iii) *that*

(3.5.4) $$\sum c_n(x) \to 1,$$

when $x \to \infty$; are necessary and sufficient that $t(x)$ should be defined by (3.5.1) *for $x \geqslant 0$, and tend to s when $x \to \infty$, whenever $s_n \to s$.*

In this case also we call the transformation T defined by $t(x)$ regular.

The theorem may be proved by an argument like that of § 3.3. But it is a corollary of Theorem 2. For, first, the conditions ensure that $t(x) \to s$ when $x \to \infty$ through any sequence (x_m) tending to ∞, and so generally. Secondly, if condition (i) is not satisfied, then either the series $\sum c_n(x_m)$ diverges for some $x_m \geqslant 0$, or

$$\overline{\lim} \sum |c_{m,n}| = \overline{\lim} \sum |c_n(x_m)| = \infty$$

when $x \to \infty$ through some sequence (x_m) tending to ∞. But then, by Theorem 2, there are sequences (s_n) for which s_n tends to s and

$$t(x_m) = \sum c_{m,n}\, s_n$$

is either not defined for some x_m or does not tend to s. This proves the necessity of (3.5.2), and that of (3.5.3) and (3.5.4) is obvious.

There are obviously similar theorems in which x tends to a finite limit a (or $a+0$ or $a-0$). These are derivable by trivial transformations, and we shall regard them as included in Theorem 5. There is also an analogue of Theorem 4 with a continuous parameter x, which we do not state explicitly.

(3) There are similar theorems concerning integral transformations

$$t(x) = \int c(x,y)s(y)\,dy;\dagger \tag{3.5.5}$$

but they are a little less symmetrical, since the kernel $c(x,y)$ may behave in a more complex way for finite x and y than a function of integral variables. We therefore confine ourselves here to the statement of *sufficient* conditions (which are all that we shall actually need), and suppose $s(y)$ bounded for all y.

THEOREM 6. *In order that*

$$s(y) \to s \quad (y \to \infty) \tag{3.5.6}$$

should imply

$$t(x) \to s \quad (x \to \infty) \tag{3.5.7}$$

for every bounded $s(y)$, it is sufficient that

$$\int |c(x,y)|\,dy < H, \tag{3.5.8}$$

where H is independent of x, that

$$\int_0^Y |c(x,y)|\,dy \to 0 \tag{3.5.9}$$

when $x \to \infty$, for every finite Y, and that

$$\int c(x,y)\,dy \to 1 \tag{3.5.10}$$

when $x \to \infty$.

The proof is like the sufficiency proof in § 3.3. We suppose first that $s = 0$. Then

$$t(x) = \int_0^Y c(x,y)s(y)\,dy + \int_Y^\infty c(x,y)s(y)\,dy = U+V,$$

† Integrations are over $(0,\infty)$ unless the contrary is indicated. The integral (3.5.5) is defined, in general, as

$$\lim_{Y\to\infty} \int_0^Y c(x,y)s(y)\,dy;$$

but in Theorem 6 it is absolutely convergent.

say. We can choose Y so that $|s(y)| < \epsilon/2H$ for $y \geqslant Y$, when

$$|V| \leqslant \frac{\epsilon}{2H} \int\limits_{Y}^{\infty} |c(x,y)|\,dy \leqslant \tfrac{1}{2}\epsilon;$$

and $U \to 0$ when Y is fixed and $x \to \infty$. Hence $t(x) \to 0$. We then pass to the general case by replacing $s(y)$ by $s_1(y) = s(y) - s$.

The transformation (3.5.5) includes those considered before as special cases. If $s(y) = s_n$, $c(x,y) = c_n(x)$ for $n \leqslant y < n+1$, then $t(x) = \sum c_n(x) s_n$, the transformation of Theorem 5. If we then restrict x to integral values m, we obtain that of Theorem 2.

The form of (3.5.9) is not quite parallel to that of (3.2.4) with $\delta_n = 0$. The parallelism would be restored if we wrote the latter condition, as we might, in the form

$$\sum_{\nu=0}^{n} |c_{m,\nu}| \to 0.$$

(4) We may also frame theorems in terms of series instead of sequences. There are two in particular which are familiar in elementary analysis,† and concern transformations of the classes $\mathfrak{T}_c$ and $\mathfrak{T}_c^*$.

THEOREM 7. *In order that $\sum \chi_n a_n$ should be convergent whenever $\sum a_n$ is convergent, it is necessary and sufficient that*

$$\sum |\Delta\chi_n| = \sum |\chi_n - \chi_{n+1}| < \infty. \tag{3.5.11}$$

THEOREM 8. *In order that $\sum \chi_n a_n$ should be convergent whenever $s_n = a_0 + a_1 + \ldots + a_n$ is bounded, it is necessary and sufficient that* (3.5.11) *should be satisfied and that χ_n should tend to zero.*

If t_m is the partial sum of $\sum \chi_n a_n$, then

$$t_m = \sum_0^m \chi_n a_n = \sum_0^{m-1} (\chi_n - \chi_{n+1}) s_n + \chi_m s_m,$$

so that

$$c_{m,n} = \Delta\chi_n \quad (0 \leqslant n < m), \qquad \chi_m \quad (n = m), \qquad 0 \quad (n > m),$$

and

$$\gamma_m = \sum_0^{m-1} |\Delta\chi_n| + |\chi_m|. \tag{3.5.12}$$

We have to show that, for this t_m, the conditions of Theorems 7 and 8 reduce to those of Theorems 1 and 3 respectively.

It is plain from (3.5.12) that (3.2.3) implies (3.5.11). Conversely, if (3.5.11) is satisfied then $\sum (\chi_n - \chi_{n+1})$ is convergent, so that χ_n tends to a limit χ. *A fortiori* it is bounded, and then (3.2.3) follows from (3.5.12). Thus (3.5.11) is equivalent to (3.2.3).

† So far as the sufficiency of their conditions is concerned.

Next, $c_{m,n} = \Delta\chi_n$ for $m > n$, so that $c_{m,n} \to \Delta\chi_n = \delta_n$ when $m \to \infty$; and

$$c_m = \sum_0^{m-1} (\chi_n - \chi_{n+1}) + \chi_m = \chi_0 = \delta.$$

Thus the conditions (3.2.4) and (3.2.5) are satisfied without further restriction on χ_n. This proves Theorem 7.

The additional condition for Theorem 8 is, by Theorem 3, that $\sum |c_{m,n}|$ should be uniformly convergent, and this, as we saw in § 3.4, is equivalent to

$$\sum |c_{m,n}| \to \sum |\delta_n|.$$

But here this is
$$\sum_0^{m-1} |\Delta\chi_n| + |\chi_m| \to \sum_0^{\infty} |\Delta\chi_n|,$$

i.e. $\sum |\Delta\chi_n| < \infty$ together with $\chi_n \to 0$. This completes the proof of Theorem 8.

We can naturally prove Theorems 7 and 8 directly without appealing to the more difficult theorems from which we have deduced them here.

(5) We conclude this section with the observation that the classes $\mathfrak{T}_r$ and $\mathfrak{T}_c^*$, both subclasses of $\mathfrak{T}_c$, are mutually exclusive. If T belongs to $\mathfrak{T}_c^*$ then $\sum |c_{m,n}|$, and *a fortiori* $\sum c_{m,n}$, is uniformly convergent, so that

$$\sum \delta_n = \sum \lim c_{m,n} = \lim \sum c_{m,n} = \lim c_m = \delta.$$

But this is impossible when T is regular, since then $\delta_n = 0$ for all n and $\delta = 1$.

3.6. Positive transformations. In this section we shall be concerned exclusively with regular transformations. There is one particularly important subclass of such transformations, in which

$$c_{m,n} \geqslant 0 \tag{3.6.1}$$

for all m, n or at any rate for $n \geqslant n_0$. We call such a transformation a *positive* (regular) transformation.

If T is regular then $c_{m,n} \to 0$, when $m \to \infty$, for every n, so that the $c_{m,n}$ with $n < n_0$ do not affect the behaviour of t_m for large m. It is therefore of little importance whether we suppose $c_{m,n} \geqslant 0$ for all m and n or only for $n \geqslant n_0$.

THEOREM 9. *If* T *is regular and positive, and s_n real, then*

$$\varliminf_{n\to\infty} s_n \leqslant \varliminf_{m\to\infty} t_m \leqslant \varlimsup_{m\to\infty} t_m \leqslant \varlimsup_{n\to\infty} s_n \tag{3.6.2}$$

for any (s_n). In particular $s_n \to s$ implies $t_m \to s$ for finite or infinite s.

If $\varliminf s_n = \sigma$ is finite, then $s_n > \sigma - \epsilon$ for $n \geqslant N = N(\epsilon)$. We may suppose $N \geqslant n_0$, and then, by (3.6.1), either $t_m = \infty$ or

$$t_m = \sum_0^N c_{m,n} s_n + \sum_{N+1}^\infty c_{m,n} s_n \geqslant \sum_0^N c_{m,n} s_n + (\sigma - \epsilon) \sum_{N+1}^\infty c_{m,n}.$$

The first term on the right tends to 0 when $m \to \infty$, and the second to $\sigma - \epsilon$, so that $t_m > \sigma - 2\epsilon$ for sufficiently large m. Hence

$$\varliminf t_m \geqslant \sigma = \varliminf s_n.$$

The proof of the last inequality (3.6.2), when $\varlimsup s_n$ is finite, is similar.

If $\varliminf s_n = \infty$ (so that $s_n \to \infty$), then $s_n > G$ for any G and $n \geqslant N = N(G) \geqslant n_0$; and either $t_m = \infty$ or $t_m > \frac{1}{2}G$ for sufficiently large m, so that $t_m \to \infty$. The case in which $\varlimsup s_n = -\infty$ is similar.

The last clause of Theorem 9 suggests a further interesting problem concerning real transformations. We may say, as in § 1.4, that a real transformation T is *totally regular* if $s_n \to s$ implies $t_m \to s$ for all *finite or infinite* s. The conditions of Theorem 2 must then be satisfied, and it is natural to ask for additional conditions necessary and sufficient for total regularity. The general conditions are rather complex, and we confine our attention to 'triangular' transformations

$$t_m = \sum_0^m c_{m,n} s_n \tag{3.6.3}$$

in which $c_{m,n} = 0$ for $n > m$.

THEOREM 10. *In order that a real transformation* (3.6.3) *should be totally regular, it is necessary and sufficient that it should be regular and positive.*

After Theorem 9, we have only to prove that, if T is totally regular, then $c_{m,n} \geqslant 0$ for $n \geqslant n_0$.

If the condition is not satisfied, there are negative $c_{m,n}$ with arbitrarily large n and, since $c_{m,n} = 0$ when $n > m$, also with arbitrarily large m. There is therefore a sequence (m_i) of m such that (1) $c_{m_i,n} < 0$ for some n, (2) if n_{m_i} is the rank of the last such $c_{m_i,n}$, then $n_{m_i} \leqslant m_i$ and n_{m_i} tends to infinity with m_i.

In what follows we consider only values of m in (m_i), and write m simply for m_i. Starting with an arbitrary m_1, we define sequences (m_r)† and (s_n) as follows. Suppose that we have determined $m_1, m_2, \ldots, m_r$, the corresponding values of n_m, and those of s_n for $n \leqslant m_r$. Since

† (m_r) is naturally a subsequence of what we first called (m_i).

$n_m \leqslant m$, n_m tends to infinity with m, and $c_{m,n} \to 0$ for each n when $m \to \infty$, we can choose m_{r+1} so that $m_{r+1} \geqslant n_{m_r+1} > m_r$ and

$$\left| \sum_0^{m_r} c_{m_r+1,n}\, s_n \right| < 1.$$

This defines (m_r) by recurrence; and we then define s_n, for $m_r < n \leqslant m_{r+1}$, by

$$s_n = n \;\; (m_r < n \leqslant m_{r+1},\, n \neq n_{m_r+1}), \qquad s_n = \frac{m_{r+1}^2}{|c_{m_r+1, n_{m_r+1}}|} \; (n = n_{m_r+1}).$$

Then $s_n \to \infty$ (since $|c_{m,n}| \leqslant H$), but

$$\begin{aligned} t_{m_r+1} &= \sum_0^{m_r} c_{m_r+1,n}\, s_n + \sum_{m_r+1}^{m_{r+1}} c_{m_r+1,n}\, s_n \\ &< 1 + m_{r+1} \sum_{m_r+1}^{m_{r+1}} |c_{m_r+1,n}| - m_{r+1}^2 < 1 + H m_{r+1} - m_{r+1}^2. \end{aligned}$$

Hence $t_m \to -\infty$ when $m = m_r$ and $r \to \infty$, and T is not totally regular.

The following examples may help the reader to appreciate the various possibilities.

(i) The transformation in which $c_{m,m} = 2$, $c_{m,m+1} = -1$, and $c_{m,n} = 0$ otherwise, is regular but not totally regular. Thus if $s_n = 2^n$, $t_m = 0$ for all m. If $s_n = 3^n$, then $s_n \to \infty$ but $t_m \to -\infty$.

(ii) The transformation

$$t_m = \frac{2}{m+1}(s_0+s_1+\dots+s_{m-1}) - \frac{m-1}{m+1} s_m,$$

which is of type (3.6.3), is regular but not totally regular; for if $s_n = n+1$, then $s_n \to \infty$ but $t_m = 1$ for all m.

(iii) The transformation defined by the matrix

$$\begin{pmatrix} 1 & -2^{-1} & 2^{-2} & 2^{-3} & . & . \\ 0 & 1 & -2^{-2} & 2^{-3} & . & . \\ 0 & 0 & 1 & -2^{-3} & . & . \\ 0 & 0 & 0 & 1 & . & . \\ . & . & . & . & . & . \end{pmatrix}$$

is totally regular. The conditions of Theorem 2 are satisfied, and

$$t_m = s_m - 2^{-m-1} s_{m+1} + 2^{-m-2} s_{m+2} + 2^{-m-3} s_{m+3} + \dots.$$

If $s_n \to \infty$, then there are two possibilities. If $\sum 2^{-n} s_n$ is divergent, then $t_m = \infty$ for all m. If $\sum 2^{-n} s_n$ is convergent, then $2^{-m-1} s_{m+1} = o(1)$ and $t_m \geqslant s_m - o(1) \to \infty$.

3.7. Knopp's kernel theorem. There is an interesting generalization of Theorem 9 for complex sequences, due to Knopp. We follow Knopp in stating it for the general integral transformation (3.5.5).

We call the transformation *positive* if $c(x, y) \geqslant 0$ for all x, y.† The conditions of Theorem 6 then reduce to

$$(a) \qquad c(x) = \int c(x,y)\, dy$$

† The condition strictly parallel to that of § 3.6 would be '$c(x, y) \geqslant 0$ for $y \geqslant Y$'. We take $Y = 0$ to avoid minor complications.

is bounded and tends to 1 when $x \to \infty$, and

$$(b) \qquad \int_0^Y c(x,y)\,dy \to 0$$

when $x \to \infty$, for every finite Y. We shall suppose throughout that these conditions are satisfied, and call such a transformation *normal.*

We state our results in terms of the complex plane $w = u+iv$ with a single point $w = \infty$ at infinity. Given any set S of points w $(w \neq \infty)$, we define the *least closed convex region K including S* (the 'convex cover' of S) as follows. If there is no closed half-plane including S, then K is the whole plane, including ∞. If there are such half-planes, then K is their common part. We count ∞ in K when S is unbounded but not when it is bounded: in any case K is closed. Thus if S is a single point, K is that point; if S consists of two points, K is the straight segment joining them; if S is the real axis, K is the real axis with the point ∞; if S is the real and imaginary axes, K is the entire plane.

Suppose now that $s(y) = u+iv$ is a complex function of the real variable y, defined for $y \geqslant 0$ and bounded in any finite interval $(0, Y)$. We define $K(s, y_0)$ as the least closed convex region K including all values of $s(y)$ for $y \geqslant y_0$: thus $K(s, y_2)$ is included in $K(s, y_1)$ if $y_2 \geqslant y_1$. Finally, we define $K(s)$, the *kernel* of $s(y)$, as the common part of all $K(s, y)$; and $K(t)$, the kernel of $t(x)$, similarly.

If $s(y)$ tends to a finite limit a when $y \to \infty$, $K(s)$ is the point a. If $s(y)$ is real, $K(s)$ is the stretch $\underline{\lim}\, s(y)$, $\overline{\lim}\, s(y)$ of the real axis, together with the point ∞ if either $\underline{\lim}\, s(y) = -\infty$ or $\overline{\lim}\, s(y) = \infty$. In any case $K(s)$ cannot be empty, since it is the limit of a decreasing sequence of non-empty closed sets; but it may consist of the single point ∞.

If $K(s)$ is the single point ∞, we say that $s(y)$ *diverges to* ∞. When $s(y)$ is real, this implies that $s(y) \to \infty$ or $s(y) \to -\infty$. The definition gives an appropriate generalization of the notion of 'proper divergence' for complex functions.

We can now state Knopp's theorem.

THEOREM 11. *If the transformation* (3.5.5) *is normal, and $t(x)$ exists for $x \geqslant 0$, then $K(t)$ is included in $K(s)$.*

In particular this is true, with the obvious modifications in the definitions, for a regular and positive transformation (3.1.3).

We may assume that $K(s)$ is not the entire plane, since in that case there is nothing to prove. Thus what we have to prove is that any point w outside $K(s)$ is also outside $K(t)$. If w is outside $K(s)$, it is

outside $K(s, y_0)$ for some y_0. Thus we have to prove that *if w is outside $K(s, y_0)$, it is outside $K(t, x_0)$ for some x_0.* We must distinguish two cases.

(1) Suppose that $w \neq \infty$. We may then suppose (making a translation if necessary) that $w = 0$. Since $K(s, y_0)$ is closed, there is a point w_0 of $K(s, y_0)$ whose distance from $w = 0$ is a minimum.† We may suppose (making a rotation if necessary) that the plane is so oriented that

$$w_0 = w_0 - w = 4d > 0.$$

Then, since $K(s, y_0)$ is convex, all of its points, and *a fortiori* all points of any $K(s, y)$ with $y > y_0$, have abscissae at least $4d$. Thus $\Re s(y) > 3d$ for $y \geqslant y_0$.

Since $s(y)$ is bounded in any finite interval of values of y, there is an M such that $|s(y)| < M$ for $0 \leqslant y \leqslant y_0$. Since

$$\int_0^{y_0} c(x, y)\, dy \to 0, \qquad \int_0^{\infty} c(x, y)\, dy \to 1$$

when $x \to \infty$, we can choose x_0 so that

$$\int_0^{y_0} c(x, y)\, dy < \frac{d}{M}, \qquad \int_{y_0}^{\infty} c(x, y)\, dy > \tfrac{2}{3}$$

for $x \geqslant x_0$. It then follows that

$$\Re t(x) = \Re\left\{\int_0^{y_0} c(x, y) s(y)\, dy\right\} + \Re\left\{\int_{y_0}^{\infty} c(x, y) s(y)\, dy\right\}$$
$$> -M(dM^{-1}) + \tfrac{2}{3}.3d = -d + 2d = d$$

for $x \geqslant x_0$, and that $w = 0$ is outside $K(t, x_0)$.

(2) Suppose that $w = \infty$. In this case $K(s, y_0)$ is bounded; and $s(y)$ is bounded for $y \geqslant y_0$, and therefore for all y. Hence $|s(y)| \leqslant N$ for some N, and

$$|t(x)| \leqslant N \int c(x, y)\, dy,$$

so that $t(x)$ is bounded. Thus $w = \infty$ is outside $K(t, x_0)$ for any x_0.

This completes the proof of Theorem 11. In particular, $t(x)$ diverges to ∞ if $s(y)$ does so.

3.8. An application of Theorem 2. Any transformation (3.1.3) may be used to define a method of summation of series: if

$$s_n = a_0 + a_1 + \ldots + a_n,$$

† Actually, since $K(s, y_0)$ is convex, there is just one such point; but this is not required for the argument.

t_m is defined by (3.1.3), and $t_m \to s$, then we may say that $\sum a_n$ is summable (T) to sum s, and write

$$s_n \to s \ \text{(T)}, \qquad \sum a_n = s \ \text{(T)}.$$

We call the method regular if T is regular, so that a regular method is one which sums every convergent series to its ordinary sum.

We shall use Theorem 2 in the next chapter to prove the regularity of the methods of summation most useful in analysis. Here we apply it to the proof of a theorem, which we shall need later, about methods of less general importance.

If

$$p_n \geqslant 0, \qquad p_0 > 0, \qquad \sum p_n = \infty \tag{3.8.1}$$

(so that $P_n = p_0+p_1+\ldots+p_n \to \infty$), and

$$t_n = \frac{p_0 s_0+p_1 s_1+\ldots+p_n s_n}{p_0+p_1+\ldots+p_n} \to s \tag{3.8.2}$$

when $n \to \infty$, then we say that

$$s_n \to s \ (\overline{\mathrm{N}}, p_n).\dagger \tag{3.8.3}$$

We prove first

THEOREM 12. *The method* $(\overline{\mathrm{N}}, p_n)$ *is regular.*

Here

$$c_{m,n} = p_n/P_m \quad (n \leqslant m), \qquad c_{m,n} = 0 \quad (n > m),$$
$$\sum |c_{m,n}| = \sum c_{m,n} = 1,$$

and $c_{m,n} \to 0$ for each n. Thus the conditions of Theorem 2 are satisfied. In particular, the (C, 1) method, in which $p_n = 1$, is regular.

In what follows we suppose, to avoid minor complications, that $p_n > 0$ for all n. We prove first

THEOREM 13. *If* $p_n > 0$ *and* $s_n \to s$ $(\overline{\mathrm{N}}, p_n)$, *then*

$$s_n - s = o(P_n/p_n).$$

For

$$p_n s_n = P_n t_n - P_{n-1} t_{n-1} = s(P_n - P_{n-1}) + o(P_n) = s p_n + o(P_n).$$

In particular, $s_n \to s$ (C, 1) implies $s_n - s = o(n)$, and so $s_n = o(n)$, $a_n = o(n)$. The theorem is one of an important class which we may call 'limitation theorems'. There is a limitation theorem associated with any useful method of summation, asserting that it cannot sum *too rapidly* divergent series.

The next theorem, which is the main theorem of this section, concerns the relations between the methods corresponding to two different sequences (p_n) and (q_n).

† The reason for this notation will appear in the next chapter.

THEOREM 14. *If* $p_n > 0$, $q_n > 0$, $\sum p_n = \infty$, $\sum q_n = \infty$, *and* **either** (a)

$$q_{n+1}/q_n \leqslant p_{n+1}/p_n, \tag{3.8.4}$$

or (b)

$$p_{n+1}/p_n \leqslant q_{n+1}/q_n \tag{3.8.5}$$

and also

$$P_n/p_n \leqslant HQ_n/q_n, \tag{3.8.6}$$

then $\sum a_n = s\ (\overline{\mathrm{N}}, p_n)$ *implies* $\sum a_n = s\ (\overline{\mathrm{N}}, q_n)$.

If

$$t_m = (p_0 s_0 + p_1 s_1 + \ldots + p_m s_m)/P_m, \quad u_m = (q_0 s_0 + q_1 s_1 + \ldots + q_m s_m)/Q_m,$$

then

$$p_0 s_0 = P_0 t_0, \qquad p_m s_m = P_m t_m - P_{m-1} t_{m-1} \quad (m > 0),$$

and so

$$u_m = \frac{1}{Q_m}\left\{\frac{q_0}{p_0} P_0 t_0 + \frac{q_1}{p_1}(P_1 t_1 - P_0 t_0) + \ldots + \frac{q_m}{p_m}(P_m t_m - P_{m-1} t_{m-1})\right\}. \tag{3.8.7}$$

Thus $u_m = \sum c_{m,n} t_n$, where

$$c_{m,n} = \left(\frac{q_n}{p_n} - \frac{q_{n+1}}{p_{n+1}}\right)\frac{P_n}{Q_m} \quad (n < m), \qquad \frac{q_m P_m}{p_m Q_m} \quad (n = m), \tag{3.8.8}$$

and 0 for $n > m$. Since $Q_m \to \infty$, $c_{m,n} \to 0$ when n is fixed and $m \to \infty$. If $s_n = 1$ for all n, then $t_m = 1$ and $u_m = 1$, so that

$$\sum c_{m,n} = 1 \tag{3.8.9}$$

for every m. Hence the transformation (3.8.7) satisfies conditions (3.2.4) and (3.2.5) of Theorem 1, with $\delta_n = 0$ and $\delta = 1$.

It remains to verify that it also satisfies (3.2.3). In case (a), $c_{m,n} \geqslant 0$, $\sum |c_{m,n}| = \sum c_{m,n}$, and (3.2.3) follows from (3.8.9). In case (b), $c_{m,n} \leqslant 0$ except when $n = m$, while $c_{m,m} > 0$. Hence

$$\sum_0^\infty |c_{m,n}| = -\sum_0^{m-1} c_{m,n} + \frac{q_m P_m}{p_m Q_m},$$

$$1 = \sum_0^\infty c_{m,n} = \sum_0^{m-1} c_{m,n} + \frac{q_m P_m}{p_m Q_m},$$

so that

$$\sum |c_{m,n}| = 2\frac{q_m P_m}{p_m Q_m} - 1 \leqslant 2H - 1,$$

by (3.8.6), and (3.2.3) is satisfied with $2H-1$ for H. Thus in either case (3.8.7) is regular, and the result of the theorem follows.

Roughly, in case (a) $\sum q_n$ diverges less rapidly than $\sum p_n$, while in case (b) it diverges more rapidly, but not too much more rapidly. If

$p_n = n^{\alpha}$, $q_n = n^{\beta}$, then the first condition is satisfied if $\alpha > \beta > -1$, and the second if $\beta > \alpha > -1$ (since P_n/p_n and Q_n/q_n are each asymptotic to a multiple of n). If $p_n = 1$, $q_n = 2^n$, then $P_n/p_n \sim n$, $Q_n/q_n \to 2$, and the theorem fails.

In fact, when $\sum q_n$ diverges rapidly, the method $(\overline{N}, q_n)$ becomes trivial, in the sense that it will sum convergent series only. This is shown more precisely by the following theorem.

THEOREM 15. *If $Q_{n+1}/Q_n \geqslant 1+\delta > 1$, then $\sum a_n$ cannot be summable $(\overline{N}, q_n)$ unless it is convergent.*

For $Q_m u_m = q_0 s_0 + \dots + q_m s_m$, and so

$$s_m = (Q_m u_m - Q_{m-1} u_{m-1})/q_m = \sum c_{m,n} u_n, \tag{3.8.10}$$

where

$$c_{m,m-1} = -Q_{m-1}/q_m, \qquad c_{m,m} = Q_m/q_m, \tag{3.8.11}$$

and the remaining $c_{m,n}$ are 0. Plainly $c_{m,n} \to 0$ when $m \to \infty$, and

$$\sum c_{m,n} = (Q_m - Q_{m-1})/q_m = 1.$$

Also $q_m \geqslant \delta Q_{m-1}$, and so

$$\sum |c_{m,n}| = (Q_{m-1} + Q_m)/q_m = 2(Q_{m-1}/q_m)+1 \leqslant 2\delta^{-1}+1.$$

Hence the transformation (3.8.10), from u_n to s_m, is regular, and $s_m \to s$ whenever $u_n \to s$.

Thus the series $1-1+1-\dots$, which is summable $(\overline{N}, 1)$, i.e. $(C, 1)$, is not summable $(\overline{N}, 2^n)$. The theorem illustrates a general principle, of which we shall find many other illustrations later, that *too violent* a method of summation tends to defeat its own object by becoming 'trivial': the more delicate methods are often the more effective. Thus the means defined by $p_n = (n+1)^{-1}$, for which

$$t_m = \frac{1}{P_m}\left(s_0 + \frac{s_1}{2} + \dots + \frac{s_m}{m+1}\right) \sim \frac{1}{\log m}\left(s_0 + \frac{s_1}{2} + \dots + \frac{s_m}{m+1}\right),$$

are more effective than the $(C, 1)$ means. They sum any series summable $(C, 1)$, and also series such as $\sum n^{-1-ci}$ for which the $(C, 1)$ method fails. We shall return to these means ('logarithmic' means) in § 4.16.

3.9. Dilution of series. One simple application of Theorem 14 is to what Chapman has described as the 'dilution' of series. The convergence or divergence of a series is not affected by the insertion of zeros as extra terms: if either of the series $a_0+a_1+a_2+\dots$ and $0+0+\dots+a_0+0+\dots+a_1+0+\dots$ converges, then the other converges to the same sum. But such a change may destroy the summability of a divergent series, or change its sum. Thus the series

$$1-1+1-\dots, \qquad 1-1+0+1-1+0+1-\dots$$

are summable $(C, 1)$ to the sums $\frac{1}{2}$ and $\frac{1}{3}$ respectively.

Let us consider, for example, the relations between the (C, 1) summability of $\sum a_n$ and that of

(i) $$\sum b_n = a_0+a_1+0+0+a_2+0+0+0+0+a_3+\dots,$$

(ii) $$\sum b_n = 0+a_0+a_1+0+a_2+0+0+0+a_3+0+\dots,$$

in which a_m occurs in ranks m^2 and 2^m respectively.

(i) If $m^2 \leqslant n < (m+1)^2$ then

$$t_n = \sum_{\nu\leqslant n} b_\nu = \sum_{m^2\leqslant n} a_m = s_{[\sqrt{n}]}.$$

Hence, if $M^2 \leqslant N < (M+1)^2$, we have

$$\frac{t_0+t_1+\dots+t_N}{N+1} = \frac{s_0+3s_1+\dots+(2M-1)s_{M-1}}{N+1}+\frac{N-M^2+1}{N+1}s_M. \tag{3.9.1}$$

The left-hand side of (3.9.1) tends to s if and only if $\sum b_n = s$ (C, 1). Also $N+1 \sim M^2$, and so the first term on the right tends to s if and only if

$$\sum a_n = s\ (\overline{\mathrm{N}}, 2n+1),$$

and this, by Theorem 14, is equivalent to $\sum a_n = s$ (C, 1). Finally, if either of these hypotheses is satisfied, $s_M = o(M)$, by Theorem 13, so that the last term in (3.9.1) is $o(M.M.M^{-2}) = o(1)$. It follows that *$\sum b_n$ is summable (C, 1) to s if and only if $\sum a_n$ is summable (C, 1) to s.*

(ii) In this case a similar argument shows that the summability of $\sum b_n$ implies that of $\sum a_n$; but the converse is not true. Suppose, for example, that $a_n = (-1)^n$, when $\sum a_n$ is summable to $\frac{1}{2}$. Then it is easily verified that

$$t_0+t_1+\dots+t_n = \tfrac{1}{3}(2^{2m}-1)$$

for $2^{2m-1} \leqslant n \leqslant 2^{2m}-1$, so that the (C, 1) mean changes from about $\frac{2}{3}$ to about $\frac{1}{3}$ when n increases over this interval.

It is natural to ask what is true of the corresponding Abelian limits. We shall prove in § 4.10 that, if $a_n = (-1)^n$, the series (i) is summable (A) to $\frac{1}{2}$. We shall also prove that $x-x^a+x^{a^2}-\dots$, where a is greater than 1, does not tend to a limit when $x \to 1$, so that, in particular, the series (ii) is not summable (A).

It is easy to prove directly that

$$\sum a_n x^{n^2} \to s \rightarrow \sum a_n x^n \to s,$$

whenever $\sum a_n x^n$ is convergent for $|x| < 1$; and we shall prove more general theorems of this kind, due to M. L. Cartwright, in Appendix V.

NOTES ON CHAPTER III

§ 3.2. The most fundamental theorem, Theorem 2, is due in substance to Toeplitz, *PMF*, 22 (1911), 113–19. Toeplitz considers only 'triangular' transformations in which $c_{m,n} = 0$ for $n > m$. The extension to general transformations, which involves no difficulty of principle, was made by Steinhaus, ibid. 121–34.

Theorem 1 was proved for triangular transformations by Kojima, *TMJ*, 12 (1917), 291–326, and independently, for general transformations, by Schur, *JM*, 151 (1921), 79–111. Schur also proved Theorem 3 in the same paper.

A number of other general theorems will be found in Dienes, ch. 12.

§ 3.4. For Dini's theorem, in its usual form, and connected theorems concerning uniform convergence, see Dini, *Grundlagen für eine Theorie der Funktionen einer veränderlichen reellen Grösse*, 148–50; Bromwich, 138–41; Hardy, *PCPS*, 19 (1918), 148–56.

We have supposed in the text that the series for t_m converges for all m. We may if we please allow it to diverge for a finite number of values of m, i.e. suppose it convergent only for $m > m_0$. Here m_0 may *prima facie* be $m_0(s)$, i.e. depend upon the sequence (s_n); but it follows from a theorem of Agnew, *BAMS*, 45 (1939), 689–730, that if the series converges for $m > m_0(s)$ whenever s_n tends to a limit, then $\sum |c_{m,n}| < \infty$ for $m > m_1$, so that we may replace $m_0(s)$ by a number m_1 independent of (s_n). See also Rogers, *JLMS*, 21 (1946), 123–8, and the note on § 3.6.

§ 3.5 (3). It may be advisable to add a note about *necessary and sufficient* conditions for the regularity of the transformation (3.5.5), though the further points at issue, depending as they do on the behaviour of $c(x,y)$, $s(y)$, and $t(x)$ for finite x and y, belong to the theory of functions of a real variable rather than to that of divergent series and integrals. There is a much fuller discussion of them in Agnew, l.c. *supra*. The materials required for the discussion will be found in Hobson, 2, ch. 7, and are due in part to Lebesgue and in part to Hobson himself.

In the text we assume $s(y)$ bounded for all y, and prove that the conditions

$$\text{(A)}\quad \gamma(x) = \int |c(x,y)|\,dy < H, \qquad \text{(B)}\quad \int c(x,y)\,dy \to 1,$$

$$\text{(C)}\quad \int_0^Y |c(x,y)|\,dy \to 0 \quad \textit{for every finite } Y$$

are sufficient. It is plain, since $t(x)$ need exist only for large x, that we may replace (A) by

$$\text{(A}')\qquad \gamma(x) < H \quad \textit{for sufficiently large } x.$$

It may be proved that (A′), (B), and

$$\text{(C}')\qquad \int_0^Y c(x,y)\,dy \to 0 \quad \textit{for every finite } Y,$$

are necessary conditions. The argument is much like that of § 3.3 (in the case $\delta_n = 0, \delta = 1$), but, as Dr. Bosanquet has pointed out to me, an additional lemma is needed, viz. *if*

$$\phi(x,Y) = \int_0^Y c(x,y)s(y)\,dy$$

exists for every finite Y and bounded $s(y)$, and $\phi(x,Y) \to 0$ when $x \to \infty$, then

$$\int_0^Y |c(x,y)|\,dy < K(Y),$$

where $K(Y)$ depends only on Y, for sufficiently large x. This result, which is true also if $s(y)$ is restricted to be continuous, is a corollary of what is proved in Hobson, 2, 432 and 441–3.

This, however, leaves a gap between (C′) and the stronger condition (C). The gap disappears when $c(x,y) \geqslant 0$; and in the general case we may fill it as follows. If we consider any bounded measurable set E of positive y, and take $s(y) = 1$ in E and 0 outside E, then $s(y) \to 0$; and therefore, if the transformation is regular,

$$\text{(C}'')\qquad \int_E c(x,y)\,dy \to 0 \quad \textit{for every bounded measurable } E.$$

This necessary condition is stronger than (C′) but weaker than (C), and it can

be shown that, with (A′) and (B), it is also sufficient. For, by another theorem of Hobson and Lebesgue, (A′) and (C″) imply

$$(\alpha) \qquad \int_0^Y c(x,y)s(y)\,dy \to 0$$

for every finite Y and bounded $s(y)$; and this is all that is needed to complete the proof of sufficiency. The theorem required, which is one of the cases of Hobson's 'general convergence theorem', will be found in Hobson, 2, 431.

Thus (A′), (B), *and* (C″) *are necessary and sufficient conditions for the regularity of* (3.5.5), when we restrict ourselves, as in the text, to bounded $s(y)$. This is proved by Hill, *BAMS*, 42 (1936), 225–8. Since we are concerned primarily with the behaviour of $s(y)$ when $y \to \infty$, there is no real loss in the restriction.

If we restrict $s(y)$ a little more, we can replace (C″) by the weaker condition (C′). Let us suppose, for example, that the only discontinuities of $s(y)$ are jumps. Then, by another case of Hobson's convergence theorem (p. 432), (A′) and (C′) imply (α), so that (A′), (B), and (C′) are necessary and sufficient when $s(y)$ is restricted in this way.

We may also make (C′) one of a necessary and sufficient set of conditions by restricting $c(x,y)$ instead of $s(y)$. If, for example, $c(x,y)$ is bounded, then by a further case of Hobson's convergence theorem (p. 423), (C′) alone implies (α), and (A′), (B), and (C′) are again necessary and sufficient for regularity. In this case $s(y)$ need not be bounded.

Finally, as Dr. Bosanquet has also pointed out to me, we may get rid of all these restrictions on either $s(y)$ or $c(x,y)$ by using yet another theorem of Hobson and Lebesgue (Hobson, 2, 422–3 and 438–41), and adding a fourth condition to (A′), (B), and (C′), viz.

(D) *if $C(x,Y)$ is the essential upper bound of $|c(x,y)|$ in $(0,Y)$, i.e. the upper bound when sets of measure zero are neglected, then $C(x,Y) < L(Y)$ for every finite Y and sufficiently large x.*

In fact (A′), (B), (C′), and (D) are necessary and sufficient conditions that $t(x) \to s$ whenever $s(y)$ is any function of y which is integrable in every finite interval and tends to s when $y \to \infty$.

This problem was considered first by Silverman, *TAMS*, 17 (1916), 284–94, and Kojima, *TMJ*, 14 (1918), 64–79 and 18 (1920), 37–45. Kojima proves an analogue of the more general Theorem 1. Both Silverman and Kojima suppose $s(y)$ bounded and restrict $c(x,y)$ more severely, assuming it continuous, uniformly in x, in any finite $(0,Y)$. This assumption enables them to replace (C′) by the much more drastic condition

(C‴) $\qquad c(x,y) \to 0$ *uniformly in any finite* $(0,Y)$:

a condition stronger even than (C).

(4) The sufficiency parts of Theorems 7 and 8 are classical and will be found in all the text-books: see, for example, Bromwich, 58–60; Hardy, 379–80. The necessity of the conditions was first proved by Hadamard, *AM*, 27 (1903), 177–83. There are, of course, corresponding theorems for integrals.

Similar theorems for double series were proved by Hardy, *PCPS*, 19 (1917), 86–95, and Kojima, *TMJ*, 17 (1920), 213–20. All these theorems have been generalized widely in different directions: see Moore, *Convergence factors*, and Ch. VI.

§ 3.6. Theorem 10 was proved by W. A. Hurwitz, *PLMS* (2), 26 (1926), 231–48. The more complex conditions for total regularity of the general transformation (3.1.3) were found by H. Hurwitz, *BAMS*, 46 (1940), 833–7.

The definition of total regularity is to be understood in a sense like that explained for the A method in § 1.4. If, for example, $s_n \to \infty$, then the series $t_m = \sum c_{m,n} s_n$ must, for each $m > m_0$, where $m_0 = m_0(s)$ may depend on the sequence (s_n) in question, either converge or diverge to ∞; and the values of t_m, when the series is convergent, must tend to ∞ with m. H. Hurwitz shows that then $c_{m,n} \geqslant 0$ for $m > m_1$ and $n > N(m)$, i.e. that there can be at most a finite number of negative coefficients in any sufficiently advanced row of the matrix of T; but this condition is (as an addition to those of Theorem 2) necessary only and not sufficient. Incidentally it follows that $m_0(s)$ may be replaced by a number m_1 independent of (s_n).

§ 3.7. Knopp, *MZ*, 31 (1930), 97–127 and 276–305.

§ 3.8. Theorem 14 is due to Cesàro, *Atti d. R. Accad. d. Lincei* [*Rendiconti* (4), 4 (1888), 452–7]. It was rediscovered by Hardy, *QJM*, 38 (1907), 269–88 (271), and is attributed to Hardy in Borel's book (p. 115). See also Bromwich, 427.

The condition $\sum p_n = \infty$ is not used explicitly in the proof, and is in fact implied by the other conditions. If condition (*a*) is satisfied, then $\sum p_n$ obviously diverges at least as rapidly as $\sum q_n$. If conditions (*b*) are satisfied, then the divergence of $\sum q_n$ implies that of $\sum (q_n/Q_n)$ by a familiar theorem of Abel [see, for example, Hardy, 421, 442]; and this, by (3.8.6), implies the divergence of $\sum (p_n/P_n)$ and so of $\sum p_n$.

§ 3.9. If $\sum a_n x^n$ is convergent for $|x| < 1$, and

$$\phi(y) = \sum a_n e^{-ny}, \qquad \psi(y) = \sum a_n e^{-n^2 y} \quad (y > 0),$$

then it follows from the formula

$$e^{-2ny} = \frac{2}{\sqrt{\pi}} \int e^{-n^2 y^2 t^2 - 1/t^2} \frac{dt}{t^2}$$

that

$$\phi(2y) = \frac{2}{\sqrt{\pi}} \int \psi(y^2 t^2) e^{-1/t^2} \frac{dt}{t^2},$$

and the theorem stated is an easy deduction. Compare the proofs of Theorems 28 and 30 (§ 4.8).

IV

SPECIAL METHODS OF SUMMATION

4.1. Nörlund means. Our main object in this chapter is to enumerate some of the methods of summation which have proved most useful in analysis and to establish their regularity by means of Theorem 2; but we shall add a good deal of additional matter. Some of the most important methods, for example Cesàro's, will be considered in much greater detail in later chapters, and these we shall dismiss shortly here.

The (C, 1) method of § 1.3 is the simplest of what are usually called Nörlund methods, though a definition substantially the same as Nörlund's had been given previously by Voronoi.

We suppose that

$$p_n \geqslant 0, \qquad p_0 > 0,\dagger \qquad P_n = p_0+p_1+\dots+p_n, \tag{4.1.1}$$

and define t_m by

$$t_m = N_m^{(p)}(s) = \frac{p_m s_0+p_{m-1}s_1+\dots+p_0 s_m}{p_0+p_1+\dots+p_m}. \tag{4.1.2}$$

If $t_m \to s$ when $m \to \infty$, and $s_n = a_0+a_1+\dots+a_n$, we shall write

$$s_n \to s, \qquad \sum a_n = s \ (\mathrm{N}, p_n). \tag{4.1.3}$$

If $p_n = 1$ for all n, then t_m is the (C, 1) mean of s_n; if

$$p_n = \binom{n+k-1}{k-1} = \frac{\Gamma(n+k)}{\Gamma(n+1)\Gamma(k)},$$

where $k > 0$, then it is the (C, k) mean.‡ Usually, as in these cases, $\sum p_n$ will be divergent, but this is not essential. Thus, if $p_0 = p_1 = 1$, and the remaining p_n are 0, then

$$t_m = \tfrac{1}{2}(s_{m-1}+s_m),$$

and we obtain the means $s_n^{(1)}$ referred to on p. 21.

4.2. Regularity and consistency of Nörlund means. We begin by determining the conditions that the means (4.1.2) should be regular.

THEOREM 16. *The condition*

$$p_n/P_n \to 0 \tag{4.2.1}$$

is necessary and sufficient for the regularity of the (N, p_n) *method.*

† This last condition is convenient, though not essential. If, e.g., $p_0 = 0$, $p_1 > 0$, and we write $p_n = q_{n-1}$, $t_m = u_{m-1}$, then u_m is an (N, q_n) mean of s_n with $q_0 > 0$.

‡ See § 5.5.

For, if $t_m = \sum c_{m,n} s_n$, then

$$c_{m,n} = p_{m-n}/P_m \quad (n \leqslant m), \qquad c_{m,n} = 0 \quad (n > m),$$

$c_{m,n} \geqslant 0$ and $\sum |c_{m,n}| = \sum c_{m,n} = 1$. Thus the first and third conditions of Theorem 2 are satisfied in any case. The second is that $c_{m,n} \to 0$ when n is fixed and $m \to \infty$. Taking $n = 0$ we obtain $p_m/P_m \to 0$, so that the condition (4.2.1) is necessary; and, since $c_{m,n} \leqslant p_{m-n}/P_{m-n}$, it is also sufficient.

We say that two methods P and Q are *consistent* if $s_n \to s$ (P), $s_n \to s'$ (Q) imply $s' = s$, i.e. if they cannot sum the same series to different sums.†

THEOREM 17. *Any two regular Nörlund methods* (N, p_n) *and* (N, q_n) *are consistent: if* $s_n \to s$ (N, p_n) *and* $s_n \to s'$ (N, q_n), *then* $s' = s$.

We write $r_n = p_0 q_n + p_1 q_{n-1} + \ldots + p_n q_0$. Then

$$N_m^{(r)}(s) = \frac{p_0 q_0 s_m + (p_0 q_1 + p_1 q_0) s_{m-1} + \ldots + (p_0 q_m + \ldots + p_m q_0) s_0}{p_0 q_0 + (p_0 q_1 + p_1 q_0) + \ldots + (p_0 q_m + \ldots + p_m q_0)}$$

$$= \frac{p_0(q_0 s_m + \ldots + q_m s_0) + \ldots + p_{m-1}(q_0 s_1 + q_1 s_0) + p_m q_0 s_0}{p_0(q_0 + \ldots + q_m) + \ldots + p_{m-1}(q_0 + q_1) + p_m q_0}$$

$$= \frac{p_0 Q_m N_m^{(q)}(s) + \ldots + p_m Q_0 N_0^{(q)}(s)}{p_0 Q_m + \ldots + p_m Q_0} = \sum_n \gamma_{m,n} N_n^{(q)}(s),$$

where $\gamma_{m,n} = p_{m-n} Q_n \Big/ \Big(\sum_{\nu=0}^{m} p_{m-\nu} Q_\nu\Big)$ if $n \leqslant m$ and $\gamma_{m,n} = 0$ if $n > m$. Here $\gamma_{m,n} \geqslant 0$, $\sum |\gamma_{m,n}| = \sum \gamma_{m,n} = 1$, and

$$\gamma_{m,n} \leqslant \frac{p_{m-n} Q_n}{(p_{m-n} + p_{m-n-1} + \ldots + p_0) q_0} = \frac{p_{m-n}}{P_{m-n}} \frac{Q_n}{q_0} \to 0$$

when $m \to \infty$, so that the means with coefficients $\gamma_{m,n}$ are regular. Hence $s_n \to s'$ (N, q_n) implies $s_n \to s'$ (N, r_n). Similarly, $s_n \to s$ (N, p_n) implies $s_n \to s$ (N, r_n); and therefore, when both hypotheses are satisfied, s and s' must be the same.

There is an interesting alternative proof which embodies an important principle, and which depends upon

THEOREM 18. *If* (N, p_n) *is regular, and* $\sum a_n = s$ (N, p_n), *then the series* $\sum a_n x^n$ *has a positive radius of convergence, and defines an analytic function* $a(x)$ *which is regular for* $0 \leqslant x < 1$ *and tends to* s *when* $x \to 1$ *through real values less than* 1.

We write

$$p(x) = \sum p_n x^n, \qquad P(x) = \sum P_n x^n, \qquad T(x) = \sum P_n t_n x^n,$$

where t_n is defined by (4.1.2), with $s_n = a_0 + a_1 + \ldots + a_n$. Since $p_n/P_n \to 0$, i.e. $P_{n-1}/P_n \to 1$, $P(x)$ is convergent for $|x| < 1$, and $p(x)$ also converges, to $(1-x)P(x)$.

† This is a much weaker assertion than that of *equivalence* (§ 4.3).

Since t_n is bounded, $T(x)$ also converges for $|x| < 1$. Since $p_0 > 0$ and $p_n \geqslant 0$, $p(x) > 0$ and $P(x) > 0$ for $0 \leqslant x < 1$.

The function $T(x)/p(x)$ is regular at the origin, and expansible in a power series $\varpi(x) = \sum \varpi_n x^n$ convergent for small x. Since $T(x) = p(x)\varpi(x)$,

$$P_n t_n = p_0 \varpi_n + p_1 \varpi_{n-1} + \ldots + p_n \varpi_0$$

for all n. But
$$P_n t_n = p_0 s_n + p_1 s_{n-1} + \ldots + p_n s_0$$

for all n, and therefore $\varpi_n = s_n$. Hence $\sum s_n x^n$ and $\sum a_n x^n$ are regular at the origin. Also

$$a(x) = \sum a_n x^n = (1-x)\sum s_n x^n = (1-x)\frac{T(x)}{p(x)} = \frac{T(x)}{P(x)},$$

and $T(x)$ and $P(x)$ are regular for $|x| < 1$. Hence $a(x)$ is regular for $|x| < 1$, except for possible poles, none of which is on the line $(0, 1)$.

Finally,
$$a(x) = \frac{T(x)}{P(x)} = \frac{\sum P_n t_n x^n}{P(x)} = \sum c_n(x) t_n,$$

where $c_n(x) = P_n x^n/P(x)$. This is a transformation from t_n to $a(x)$, which plainly satisfies the conditions of Theorem 5.† Thus $t_m \to s$ implies $a(x) \to s$, and this completes the proof of the theorem.

Theorem 17 is a corollary, since the sum of $\sum a_n$, if it exists, does not depend on the special values of p_n.

Theorem 18 may be regarded as 'Abel's theorem' for a regular Nörlund method. We cannot say that $\sum a_n = s$ (N, p_n) implies $s_n \to s$ (A), since $\sum a_n x^n$ will not usually converge for $0 < x < 1$; but the Abelian limit exists in a generalized sense. We may also regard the theorem as embodying a 'limitation theorem', viz. $a_n = O(e^{cn})$ for some c.

4.3. Inclusion. We now consider questions of *inclusion* and *equivalence*. We say that Q includes P if $s_n \to s$ (P) implies $s_n \to s$ (Q), and that the methods are equivalent if each includes the other. If Q includes P, but is not equivalent to P, then we shall say that Q is *stronger* than P. Here we are concerned with the case in which P is (N, p_n) and Q is (N, q_n).

If (N, p_n) and (N, q_n) are regular, then $p_n/P_n \to 0$ and $q_n/Q_n \to 0$, and the series

$$p(x) = \sum p_n x^n,\quad P(x) = \sum P_n x^n,\quad q(x) = \sum q_n x^n,\quad Q(x) = \sum Q_n x^n \tag{4.3.1}$$

are convergent for $|x| < 1$. The series

$$k(x) = \sum k_n x^n = q(x)/p(x) = Q(x)/P(x), \tag{4.3.2}$$

$$l(x) = \sum l_n x^n = p(x)/q(x) = P(x)/Q(x), \tag{4.3.3}$$

are convergent for small x, and

$$k_0 p_n + \ldots + k_n p_0 = q_n,\qquad k_0 P_n + \ldots + k_n P_0 = Q_n, \tag{4.3.4}$$

$$l_0 q_n + \ldots + l_n q_0 = p_n,\qquad l_0 Q_n + \ldots + l_n Q_0 = P_n. \tag{4.3.5}$$

† In the form with $0 \leqslant x < 1$, $x \to 1$: see the remark on p. 50 after the proof of Theorem 5. We shall take such variations of the theorem for granted later.

THEOREM 19. *If* (N,p_n) *and* (N,q_n) *are regular, then, in order that* (N,q_n) *should include* (N,p_n), *it is necessary and sufficient that*

$$|k_0|P_n+|k_1|P_{n-1}+\dots+|k_n|P_0 \leqslant HQ_n, \tag{4.3.6}$$

where H is independent of n, and that

$$k_n/Q_n \to 0. \tag{4.3.7}$$

If $P_n \to \infty$, the second condition may be omitted.

If $s(x) = \sum s_n x^n$, then

$$\sum Q_n N_n^{(q)}(s)x^n = \sum (q_0 s_n+\dots+q_n s_0)x^n = q(x)s(x)$$

for small x, and similarly $\sum P_n N_n^{(p)}(s) = p(x)s(x)$. Hence

$$\sum Q_n N_n^{(q)}(s) = \sum k_n x^n \sum P_n N_n^{(p)}(s)x^n,$$

$$Q_n N_n^{(q)}(s) = k_n P_0 N_0^{(p)}(s)+k_{n-1}P_1 N_1^{(p)}(s)+\dots+k_0 P_n N_n^{(p)}(s).$$

Thus
$$N_n^{(q)}(s) = \sum c_{n,r} N_r^{(p)}(s),$$

where $c_{n,r}$ is $k_{n-r}P_r/Q_n$ if $r \leqslant n$ and 0 if $r > n$. The first condition of Theorem 2 is (4.3.6). The third is satisfied automatically because of (4.3.4). Finally, $Q_{n-r} \sim Q_n$, for any fixed r, by Theorem 16, when $n \to \infty$, and the second condition reduces to $k_{n-r}/Q_{n-r} \to 0$, which is (4.3.7).

If $P_n \to \infty$ then, given G, we can choose r so that $P_r > G$. If also (4.3.6) is satisfied, then

$$G|k_{n-r}| \leqslant HQ_n, \qquad \overline{\lim}\frac{|k_{n-r}|}{Q_{n-r}} \leqslant \frac{H}{G}\lim\frac{Q_n}{Q_{n-r}} = \frac{H}{G},$$

and (4.3.7) follows from (4.3.6). Thus (4.3.7) may be discarded when $\sum p_n = \infty$.

If $p_n = 1$, $P_n = n+1$, then

$$p(x) = (1-x)^{-1}, \qquad k(x) = (1-x)q(x), \qquad k_0 = q_0, \qquad k_n = q_n-q_{n-1} \quad (n > 0),$$

and (4.3.6) becomes

$$(n+1)q_0+n|q_1-q_0|+\dots+|q_n-q_{n-1}| \leqslant HQ_n,$$

which is plainly satisfied if q_n increases with n. Thus we obtain

THEOREM 20. *If (N,q_n) is a regular Nörlund method with increasing q_n, then $s_n \to s$ $(\mathrm{C},1)$ implies $s_n \to s$ (N,q_n).*

4.4. Equivalence.

We next prove

THEOREM 21. *In order that two regular Nörlund methods (N,p_n) and (N,q_n) should be equivalent, it is necessary and sufficient that*

$$\sum |k_n| < \infty, \qquad \sum |l_n| < \infty. \tag{4.4.1}$$

(1) *The conditions are necessary.* Since $p_0 > 0$ and $q_0 > 0$, $k_0 > 0$ and $l_0 > 0$. Since (N,q_n) includes (N,p_n), it follows from Theorem 19 that $k_0 P_n \leqslant HQ_n$. Thus P_n/Q_n is bounded, and similarly Q_n/P_n is bounded.

By Theorem 19,

$$|k_0|+|k_1|\frac{P_{n-1}}{P_n}+\ldots+|k_r|\frac{P_{n-r}}{P_n} \leqslant H\frac{Q_n}{P_n}$$

for $r \leqslant n$. Fixing r, and making $n \to \infty$, we see that

$$|k_0|+|k_1|+\ldots+|k_r| \leqslant H \overline{\lim}\,(Q_n/P_n).$$

Thus $\sum |k_n| < \infty$, and similarly $\sum |l_n| < \infty$.

(2) *The conditions are sufficient.* If $\sum |k_n| < \infty$ then $k_n \to 0$ and $k_n/Q_n \to 0$. Also

$$P_n = Q_0 l_n + Q_1 l_{n-1} + \ldots + Q_n l_0 \leqslant Q_n \sum |l_n|,$$

$$P_n|k_0|+P_{n-1}|k_1|+\ldots+P_0|k_n| \leqslant Q_n \sum |k_n| \sum |l_n|.$$

Thus the conditions imply those of Theorem 19, with $H = \sum |k_n| \sum |l_n|$, and (N, q_n) includes (N, p_n). Similarly (N, p_n) includes (N, q_n).

It is plain that the conditions cannot be satisfied when $p(x)$ and $q(x)$ are rational and one of them has a zero, inside or on the unit circle, which is not a zero of the other. If, for example, $p_n = 2n+1$, $q_n = n+1$, then

$$p(x) = \frac{1+x}{(1-x)^2}, \qquad q(x) = \frac{1}{(1-x)^2}, \qquad l(x) = 1+x, \qquad k(x) = \frac{1}{1+x},$$

so that $\sum |k_n| = \infty$. Also

$$|k_0|P_n+\ldots+|k_n|P_0 = P_0+\ldots+P_n$$

is of order n^3, so that (4.3.6) of Theorem 19 is not satisfied, and (N, q_n) does not include (N, p_n). We shall return to this example in § 5.16.

4.5. Another theorem concerning inclusion. We now apply Theorem 19 to the proof of a criterion for inclusion of a more special kind. We are here interested primarily in cases in which P_n tends slowly to infinity, and p_n will be a decreasing function of n.

We shall use a lemma of independent interest.

THEOREM 22. *If $p(x) = \sum p_n x^n$ is convergent for $|x| < 1$, and*

(4.5.1) $$p_0 = 1, \qquad p_n > 0, \qquad \frac{p_{n+1}}{p_n} \geqslant \frac{p_n}{p_{n-1}} \quad (n > 0),$$

then

(4.5.2) $$\{p(x)\}^{-1} = 1 - c_1 x - c_2 x^2 - \ldots$$

where $c_n \geqslant 0$, $\sum c_n \leqslant 1$. If $\sum p_n = \infty$, then $\sum c_n = 1$.

It follows from the conditions that p_{n+1}/p_n increases with n, and tends to a limit which cannot exceed 1. Hence p_n decreases with n. We suppose that

$$\{p(x)\}^{-1} = \gamma_0 + \gamma_1 x + \gamma_2 x^2 + \ldots$$

for small x. Then $\gamma_0 = 1$, and it is only necessary to prove that $c_n = -\gamma_n \geqslant 0$ for $n > 0$, the remaining clauses of the theorem being corollaries, since $p(x) > 0$ for $0 \leqslant x < 1$.

We have

$$(4.5.3) \qquad \gamma_0 p_n + \dots + \gamma_n p_0 = 0, \qquad \gamma_0 p_{n+1} + \dots + \gamma_{n+1} p_0 = 0$$

for $n > 0$. It follows from (4.5.3) that

$$p_{n+1}(\gamma_1 p_{n-1} + \dots + \gamma_n p_0) = p_n(\gamma_1 p_n + \dots + \gamma_{n+1} p_0);$$

and so that $\qquad \gamma_{n+1} = a_{1,n}\gamma_1 + a_{2,n}\gamma_2 + \dots + a_{n,n}\gamma_n,$

where

$$a_{m,n} = \frac{p_{n+1}}{p_n}\frac{p_{n-m}}{p_0} - \frac{p_{n-m+1}}{p_0} = \frac{p_{n-m}}{p_0}\left(\frac{p_{n+1}}{p_n} - \frac{p_{n-m+1}}{p_{n-m}}\right) \geqslant 0.$$

Thus if $\gamma_1, \gamma_2, \dots, \gamma_n$ have the same sign, γ_{n+1} has the same sign also. Since $\gamma_1 = -\gamma_0 p_1/p_0 = -p_1 < 0$, it follows that $\gamma_n \leqslant 0$ for $n = 1, 2, \dots$.

Incidentally it appears that $\sum \gamma_n x^n$ is in fact absolutely convergent for $|x| \leqslant 1$.

We can now prove

THEOREM 23. *If* (i) (N, p_n) *and* (N, q_n) *are regular Nörlund methods;* (ii) p_n *satisfies* (4.5.1); (iii) $q_n > 0$; *and* (iv) $p_n/p_{n-1} \leqslant q_n/q_{n-1}$ $(n > n_0)$; *then* (N, q_n) *includes* (N, p_n).

We suppose first that $n_0 = 0$, i.e. that (iv) is satisfied for all $n > 0$. Since

$$(q_0 + q_1 x + \dots)(1 - c_1 x - \dots) = k_0 + k_1 x + \dots,$$

$$(p_0 + p_1 x + \dots)(1 - c_1 x - \dots) = 1,$$

we have $k_0 = q_0$ and

$$q_n - c_1 q_{n-1} - \dots - c_n q_0 = k_n, \qquad p_n - c_1 p_{n-1} - \dots - c_n p_0 = 0$$

for $n > 0$. Hence

$$\frac{k_n}{q_n} = 1 - c_1\frac{q_{n-1}}{q_n} - \dots - c_n\frac{q_0}{q_n} \geqslant 1 - c_1\frac{p_{n-1}}{p_n} - \dots - c_n\frac{p_0}{p_n} = 0,$$

and $k_n \geqslant 0$ for all n. We can now verify at once that the conditions of Theorem 19 are satisfied. For the first

$$|k_0|P_n + \dots + |k_n|P_0 = k_0 P_n + \dots + k_n P_0 = Q_n;$$

and for the second

$$k_n p_0 \leqslant k_0 p_n + \dots + k_n p_0 = q_n,$$

so that $k_n = O(q_n) = o(Q_n)$, by Theorem 16. This proves the theorem in the special case $n_0 = 0$.

Passing to the general case, we have

$$p_n/p_{n-1} \leqslant q_n/q_{n-1} \quad (n = n_0+1,\, n_0+2,\ldots).$$

We write

$$r_n = p_n \quad (n = n_0,\, n_0+1,\ldots),$$

and increase p_{n_0-1}, if necessary, to a value r_{n_0-1} such that

$$r_{n_0}/r_{n_0-1} \leqslant r_{n_0+1}/r_{n_0}, \qquad r_{n_0}/r_{n_0-1} \leqslant q_{n_0}/q_{n_0-1};$$

then p_{n_0-2} to a value r_{n_0-2} such that

$$r_{n_0-1}/r_{n_0-2} \leqslant r_{n_0}/r_{n_0-1}, \qquad r_{n_0-1}/r_{n_0-2} \leqslant q_{n_0-1}/q_{n_0-2};$$

and so on down to p_0 and r_0. Then

$$r_n/r_{n-1} \leqslant r_{n+1}/r_n, \qquad r_n/r_{n-1} \leqslant q_n/q_{n-1}$$

for $n > 0$; and $\rho_n = r_n/r_0$ satisfies

$$\rho_0 = 1, \qquad \rho_n > 0, \qquad \rho_{n+1}/\rho_n \geqslant \rho_n/\rho_{n-1}, \qquad \rho_n/\rho_{n-1} \leqslant q_n/q_{n-1}$$

for $n > 0$. It follows from what we have proved already that (N, q_n) includes (N, ρ_n), or, what is the same thing, that (N, q_n) includes (N, r_n).

It is therefore sufficient to prove that (N, r_n) includes (N, p_n). We write

$$r_n = p_n + \delta_n \quad (n = 0, 1,\ldots, n_0-1)$$

so that

$$r(x) = \sum r_n x^n = p(x) + \sum_0^{n_0-1} \delta_n x^n = p(x) + \delta(x),$$

say. By Theorem 22,

$$\{p(x)\}^{-1} = 1 - \sum c_n x^n = \sum \gamma_n x^n,$$

where $\sum |\gamma_n| \leqslant 1 + \sum c_n \leqslant 2$. Thus, if

$$k(x) = r(x)/p(x) = \sum k_n x^n,$$

we have

$$\sum_0^\infty k_n x^n = 1 + \frac{\delta(x)}{p(x)} = 1 + \sum_0^{n_0-1} \delta_n x^n \sum_0^\infty \gamma_n x^n$$

and so

$$\sum |k_n| \leqslant 1 + \sum \delta_n \sum |\gamma_n| \leqslant 1 + 2\sum \delta_n = H,$$

say. Hence, first, $k_n = o(1) = o(R_n)$; and secondly,

$$|k_0|P_n + \ldots + |k_n|P_0 \leqslant HP_n \leqslant HR_n.$$

These are the two conditions of Theorem 19, with r for q, and therefore (N, r_n) includes (N, p_n).

4.6. Euler means. We defined $\sum a_n = s\ (\mathrm{E}, 1)$, in §1.3(4), as meaning $\sum 2^{-n-1} b_n = s$, where

$$b_n = a_0 + \binom{n}{1} a_1 + \binom{n}{2} a_2 + \ldots + a_n.$$

Here

$$t_m = \sum_{n=0}^{m} \frac{b_n}{2^{n+1}},$$

and we can express t_m in terms of s_n as follows. If E is the operator defined by $Eu_n = u_{n+1}$, then $b_n = (1+E)^n a_0$ and

$$t_m = \frac{1}{2}\sum_{n=0}^{m}\left(\frac{1+E}{2}\right)^n a_0.$$

Now

$$\frac{1}{2}\sum_{n=0}^{m}\left(\frac{1+x}{2}\right)^n = \frac{1}{2}\,\frac{1-\{\frac{1}{2}(1+x)\}^{m+1}}{1-\frac{1}{2}(1+x)} = 2^{-m-1}\frac{(1+1)^{m+1}-(1+x)^{m+1}}{1-x}$$

$$= 2^{-m-1}\sum_{n=1}^{m+1}\binom{m+1}{n}\frac{1-x^n}{1-x} = 2^{-m-1}\sum_{n=1}^{m+1}\binom{m+1}{n}(1+x+x^2+\dots+x^{n-1});$$

and, since this is an identity between polynomials, we may use it with E for x. Thus

$$t_m = 2^{-m-1}\sum_{n=1}^{m+1}\binom{m+1}{n}(1+E+\dots+E^{n-1})a_0$$

$$= 2^{-m-1}\sum_{n=1}^{m+1}\binom{m+1}{n}s_{n-1} = 2^{-m-1}\sum_{n=0}^{m}\binom{m+1}{n+1}s_n.$$

Hence $t_m = \sum c_{m,n}\,s_n$, where

$$c_{m,n} = 2^{-m-1}\binom{m+1}{n+1}\quad (n \leqslant m),\qquad c_{m,n} = 0\quad (n > m),$$

$$c_{m,n} \geqslant 0,\qquad \sum |c_{m,n}| = \sum c_{m,n} = 1-2^{-m-1} \to 1,$$

and $c_{m,n} < 2^{-m-1}(m+1)^{n+1} \to 0$ when $m \to \infty$. Thus the conditions of Theorem 2 are satisfied, and

THEOREM 24. *The* (E, 1) *method is regular.*

4.7. Abelian means. If

(4.7.1) $$0 \leqslant \lambda_0 < \lambda_1 < \lambda_2 < \dots,\qquad \lambda_n \to \infty,$$

$\sum a_n e^{-\lambda_n x}$ is convergent for all positive x, and

(4.7.2) $$f(x) = \sum a_n e^{-\lambda_n x} \to s$$

when $x \to 0$, then we say that $\sum a_n$ is summable (A, λ_n), or (A, λ), to sum s, and write

(4.7.3) $$\sum a_n = s\ (\mathrm{A}, \lambda).$$

When $\lambda_n = n$, the (A, λ) method is the A method of § 1.3 (2). We shall sometimes write (A, k) for (A, n^k).

It will be convenient to consider a more general method of summation. We suppose that $(\phi_n(x))$ is a sequence of functions defined in an interval $0 < x \leqslant X$, and that

(4.7.4) $$\phi_n(x) \to 1,$$

for each n, when $x \to 0$. If

(4.7.5) $$\phi(x) = \sum a_n \phi_n(x)$$

is convergent in some interval $0 < x \leqslant X_1 \leqslant X$, and $\phi(x) \to s$ when $x \to 0$, we say that $\sum a_n$ is summable (ϕ) to sum s.

THEOREM 25. *In order that the 'ϕ' method should be regular, it is necessary and sufficient that*

(4.7.6) $$\sum |\phi_n(x) - \phi_{n+1}(x)| < H,$$

where H is independent of x, in some interval $0 < x \leqslant \xi$. In particular this condition is satisfied if

(4.7.7) $$0 \leqslant \phi_{n+1}(x) \leqslant \phi_n(x).$$

(1) *The condition is sufficient.* It follows from (4.7.4) that $|\phi_0(x)| < H_1$, say, in some interval $(0, \xi)$, and from (4.7.6) that

$$|\phi_n(x)| \leqslant |\phi_0(x)| + \sum_0^{n-1} |\phi_{\nu+1}(x) - \phi_\nu(x)| < H + H_1$$

in some such interval. Thus the system (ϕ_n) is uniformly bounded in such an interval.

Suppose first that $s_n \to 0$. Then

(4.7.8) $$\sum_0^N a_n \phi_n = \sum_0^{N-1} s_n(\phi_n - \phi_{n+1}) + s_N \phi_N.$$

The last term tends to 0, and so

(4.7.9) $$\phi(x) = \sum s_n\{\phi_n(x) - \phi_{n+1}(x)\} = \sum c_n(x) s_n,$$

say. It follows from (4.7.4) that $c_n(x) \to 0$, for each n, when $x \to 0$, and from (4.7.6) that $\sum |c_n(x)| < H$; and hence $\phi(x) \to 0$.†

If $s_n \to s$, $a_0' = a_0 - s$, and $a_n' = a_n$ for $n > 0$, then $s_n' \to 0$ and

$$\psi(x) = \sum a_n' \phi_n(x) = \phi(x) - s\phi_0(x) \to 0$$

when $x \to 0$. Hence, by (4.7.4), $\phi(x) \to s$.

(2) *The condition is necessary.* It is enough to prove that it is satisfied if $s_n \to 0$ always implies $\phi(x) \to 0$. We consider first a small fixed x. Since $\sum a_n \phi_n$ is convergent whenever $s_n \to s$, it follows from Theorem 7 that $\sum |\phi_n - \phi_{n+1}| < \infty$ for each such x. Hence $\phi_n(x)$ is bounded for such an x, and we can deduce (4.7.9) as under (1). It then follows from Theorem 5 that $\sum |c_n(x)| < H$ for small x, and this is (4.7.6).

Finally, if ϕ_n satisfies (4.7.7), then

$$\sum |\phi_n - \phi_{n+1}| = \sum (\phi_n - \phi_{n+1}) = \phi_0 - \lim \phi_n \leqslant \phi_0 < H_1,$$

since ϕ_0 is bounded in $(0, \xi)$.

† Here we appeal to the analogue of Theorem 4 mentioned on p. 50 but not stated explicitly. We cannot appeal to Theorem 5 because $\sum c_n(x)$ does not necessarily tend to 1.

There is plainly a variant of the theorem in which x is replaced by an integral parameter m which tends to infinity.

We add a remark about the special case in which ϕ_n is a positive decreasing function of n. There is a simple but useful theorem which we shall need later and which it is convenient to prove here.

THEOREM 26. *If b_n increases to infinity with n, and $\sum u_n$ is convergent, then*

$$V_n = v_0+v_1+...+v_n = b_0 u_0+b_1 u_1+...+b_n u_n = o(b_n). \tag{4.7.10}$$

If $w_n = u_n+u_{n+1}+...$, then $w_n \to 0$. Also

$$V_n = \sum_0^n b_m(w_m-w_{m+1}) = b_0 w_0+\sum_1^n (b_m-b_{m-1})w_m-b_n w_{n+1};$$

and so $V_n = T_n b_n+o(b_n)$, where

$$T_n = \frac{b_0}{b_n}w_0+\frac{b_1-b_0}{b_n}w_1+...+\frac{b_n-b_{n-1}}{b_n}w_n \to 0,$$

by Theorem 12.

It follows from Theorem 26 that (4.7.9) is true whenever ϕ_n decreases to 0 as $n\to\infty$ and $\sum a_n\phi_n$ is convergent. For, taking

$$b_n = \phi_n^{-1}, \qquad u_n = a_n\phi_n, \qquad v_n = a_n = b_n u_n,$$

in Theorem 26, we see that $s_n\phi_n \to 0$, and (4.7.9) follows from (4.7.8).

The conditions of Theorem 25 are plainly satisfied when $\phi_n(x) = e^{-\lambda_n x}$. Hence

THEOREM 27. *The* (A, λ) *method is regular. In particular, the* A *method is regular.*

There is no general theorem for Abelian methods corresponding to Theorem 17: different methods may well sum the same series to different sums. Thus $1-1+1-...$ is summable (A) to sum $\frac{1}{2}$, but summable (A,λ), when (λ_n) is the sequence 0, 1, 3, 4, 6, 7,..., to $\frac{1}{3}$: see § 3.9.

4.8. A theorem of inclusion for Abelian means. In this section we prove one theorem of inclusion for two systems of Abelian means. Others will be proved in Appendix V. As is to be expected after the last remark of § 4.7, all these theorems have a very special character.

THEOREM 28. *If* (i) $\lambda_0 \geqslant 1$, $\mu_n = \log\lambda_n$,

(ii) $\sum a_n = s\ (A,\lambda)$,

(iii) $\sum a_n e^{-\mu_n y} = \sum a_n e^{-y\log\lambda_n} = \sum a_n \lambda_n^{-y}$

is convergent for $y>0$, then $\sum a_n = s\ (A,\mu)$.

We need two preliminary theorems (the first of which is important in itself).

THEOREM 29. *Suppose that $f_0(x), f_1(x), f_2(x),\ldots$ is a sequence of functions defined in an interval of values of x; that*

(4.8.1) $|f_0(x)| < H,$ (4.8.2) $\sum |f_n(x)-f_{n+1}(x)| < K,$

where H and K are independent of x; and that $\sum b_n$ is convergent. Then $\sum b_n f_n(x)$ is uniformly convergent.

In particular the result is true if $f_n(x)$ is monotonic in n, and uniformly bounded, since then

$$\sum |f_n-f_{n+1}| = |\sum (f_n-f_{n+1})| = \left|f_0-\lim_{n\to\infty} f_n\right|.$$

We might replace the interval by any set of real or complex x.

We note first that, after (4.8.2), $\sum (f_n-f_{n+1})$ is convergent for each x, so that $f_n(x) \to f(x)$, say, when $n \to \infty$. Also

$$|f_n| \leqslant |f_0|+\sum_0^{n-1}|f_\nu-f_{\nu+1}| < H+K. \tag{4.8.3}$$

We suppose that $\sum b_n = B$, and write

$$B_n = b_0+b_1+\ldots+b_n, \qquad \beta_n = B_n-B,$$

with the convention $B_{-1} = 0$, $\beta_{-1} = -B$. Then $\beta_n \to 0$, and we can choose N_0 so that $|\beta_n| < \epsilon$ for $n \geqslant N_0-1$. Also

$$\sum_N^{N'} b_n f_n = \sum_N^{N'} (\beta_n-\beta_{n-1})f_n = -\beta_{N-1}f_N+\sum_N^{N'-1}\beta_n(f_n-f_{n+1})+\beta_{N'}f_{N'} \tag{4.8.4}$$

for $N' \geqslant N \geqslant 0$. It follows from (4.8.2)–(4.8.4) that

$$\left|\sum_N^{N'} b_n f_n\right| \leqslant 2\epsilon(H+K)+\epsilon K = (2H+3K)\epsilon$$

for $N' \geqslant N \geqslant N_0$ and each x; and this proves the theorem.

When $N = 0$ and $N' \to \infty$, (4.8.4) gives

$$\sum b_n f_n = Bf_0+\sum \beta_n(f_n-f_{n+1}) = Bf_0+\sum (B_n-B)(f_n-f_{n+1}). \tag{4.8.5}$$

Since $f_0 = f+\sum (f_n-f_{n+1})$, we have also the simpler formula

$$\sum b_n f_n = Bf+\sum B_n(f_n-f_{n+1}). \tag{4.8.6}$$

But the series on the right of (4.8.6) is *not* usually uniformly convergent. Suppose, for example, that

$$f_0 = 1, \qquad f_n = x^n \quad (n > 0,\ 0 \leqslant x \leqslant 1),$$

so that $f = 0$ for $x < 1$ and $f = 1$ for $x = 1$, that $b_0 = 1$, and that $b_n = 0$ for $n > 0$. Then $B_n = 1$, $\beta_n = 0$ for $n \geqslant 0$, and (4.8.6) becomes

$$1 = f+\sum (x^n-x^{n+1}).$$

The last series is neither uniformly convergent nor continuous, having the sum 1 for $x < 1$ (when $f = 0$) and 0 for $x = 1$ (when $f = 1$).

The conditions of the theorem are plainly satisfied when

$$f_n(x) = e^{-\lambda_n x}\ (x \geqslant 0) \quad \text{or} \quad e^{-\lambda_n(x-x_0)}\ (x \geqslant x_0).$$

If $\sum b_n$ is convergent, then $\sum b_n e^{-\lambda_n x}$ is uniformly convergent for $x \geqslant 0$; and if the last series is convergent for $x > 0$, then it is uniformly convergent in any interval $x \geqslant x_0 > 0$.

Our second preliminary theorem is

THEOREM 30. *If $\lambda_0 \geqslant 1$, $\phi(w) = \sum a_n e^{-\lambda_n w}$, and $\sum a_n \lambda_n^{-y}$ is convergent for $y > 0$, then*

$$\psi(y) = \sum a_n \lambda_n^{-y} = \frac{1}{\Gamma(y)} \int_0^\infty w^{y-1}\phi(w)\,dw. \tag{4.8.7}$$

If w and y are fixed, then $\lambda_n^y e^{-\lambda_n w}$ decreases from a certain n, so that

$$\sum a_n e^{-\lambda_n w} = \sum (\lambda_n^y e^{-\lambda_n w} . a_n \lambda_n^{-y})$$

is convergent for $w > 0$. Hence it is uniformly convergent in any interval $0 < \omega \leqslant w \leqslant W < \infty$, and

$$\int_\omega^W w^{y-1}\phi(w)\,dw = \sum a_n \int_\omega^W w^{y-1}e^{-\lambda_n w}\,dw. \tag{4.8.8}$$

We wish to replace ω and W here by 0 and ∞. For this, it is sufficient to show that the series

$$\sum a_n \int_0^\omega w^{y-1}e^{-\lambda_n w}\,dw, \qquad \sum a_n \int_W^\infty w^{y-1}e^{-\lambda_n w}\,dw \tag{4.8.9}$$

are convergent and tend to 0 when $\omega \to 0$ and $W \to \infty$.

The first series (4.8.9) is

$$\sum \frac{a_n}{\lambda_n^y} \int_0^{\lambda_n \omega} u^{y-1}e^{-u}\,du = \sum \frac{a_n}{\lambda_n^y} \chi_n(\omega) = \sum b_n \chi_n(\omega),$$

say. Here $\sum b_n$ is convergent, by hypothesis, while $\chi_n(\omega)$ is positive, increases with n, and is uniformly bounded for all n and ω. It follows from Theorem 29 that $\sum b_n \chi_n(\omega)$ converges uniformly in ω, and therefore tends to 0 when $\omega \to 0$. The proof that the second series (4.8.9) tends to 0, when $W \to \infty$, is similar.

It is now easy to prove Theorem 28. We may suppose $s = 0$, so that $\phi(w) \to 0$ when $w \to 0$. Then

$$\psi(y) = \frac{1}{\Gamma(y)}\left(\int_0^\delta + \int_\delta^\infty\right) w^{y-1}\phi(w)\,dw = P+Q,$$

say. Since $\phi(w) \to 0$, we can choose δ so that $|\phi(w)| < \epsilon$ for $0 < w \leqslant \delta$, and then

$$|P| < \frac{\epsilon}{\Gamma(y)} \int_0^\delta w^{y-1}\, dw = \frac{\epsilon\delta^y}{y\Gamma(y)} < 2\epsilon$$

for sufficiently small y. Also $\sum a_n e^{-(\lambda_n-\lambda_0)w}$ is uniformly convergent for $w \geqslant \delta$, so that $|\phi(w)| \leqslant He^{-\lambda_0 w}$, where H is independent of w. Hence

$$|Q| \leqslant \frac{H}{\delta^{1-y}\Gamma(y)} \int_\delta^\infty e^{-\lambda_0 w}\, dw \to 0$$

when $y \to 0$, and $|\psi(y)| < 3\epsilon$ for sufficiently small y.

It is easy to give examples of series summable $(\mathrm{A}, \log n)$ but not summable (A): we shall see, for example, that $\sum n^{-1-ci}$, where $c > 0$, is such a series.†

4.9. Complex methods. It is often important in applications to consider the limit of a series $\sum a_n e^{-\lambda_n z}$ when $z \to 0$ along a path in the complex plane, usually a straight line making an acute angle with the positive real axis.

If $z = x+iy$, $\sum a_n e^{-\lambda_n z}$ is convergent for $x > 0$, and

$$f(z) = \sum a_n e^{-\lambda_n z} \to s \tag{4.9.1}$$

when $z \to 0$ along any path lying in the angle $|y| \leqslant x \tan\alpha$, where $0 < \alpha < \frac{1}{2}\pi$, then we shall say that

$$\sum a_n = s\ (\mathrm{A}, \lambda, \alpha). \tag{4.9.2}$$

This method also is regular.

THEOREM 31. *If $\sum a_n$ converges to s, then $\sum a_n e^{-\lambda_n z} \to s$ when $z \to 0$, uniformly in the angle $|y| \leqslant x \tan\alpha$.*

We may suppose $s = 0$. If $x > 0$, then

$$f(z) = \sum a_n e^{-\lambda_n z} = \sum s_n(e^{-\lambda_n z} - e^{-\lambda_{n+1} z}),$$

or $f(z) = \sum c_n(z) s_n$, where

$$c_n(z) = e^{-\lambda_n z} - e^{-\lambda_{n+1} z} = \Delta e^{-\lambda_n z}.$$

Also

$$\sum |c_n(z)| = \sum |\Delta e^{-\lambda_n z}| = \sum \left| \int_{\lambda_n}^{\lambda_{n+1}} z e^{-tz}\, dt \right|$$

$$\leqslant |z| \sum \int_{\lambda_n}^{\lambda_{n+1}} e^{-tx}\, dt = \frac{|z|}{x} \sum \Delta e^{-\lambda_n x} \leqslant e^{-\lambda_0 x} \sec\alpha.$$

† See § 7.9. When we speak of summability (A, log n) we suppose our series to begin with the term in a_1, so that λ_0 is replaced by λ_1.

Hence, if we choose $N = N(\epsilon)$ so that $|s_n| < \epsilon$ for $x \geqslant N$, we have

$$\left|\sum_{N}^{N'} c_n(z)s_n\right| < \epsilon \sec\alpha$$

for $N' \geqslant N$, so that $\sum c_n(z)s_n$ is uniformly convergent. Since $c_n(z) \to 0$ when $z \to 0$, $f(z) \to 0$ uniformly in the angle.

We have proved the theorem directly: it might also be deduced from appropriate modifications of Theorem 5 or Theorem 29.

4.10. Summability of $1-1+1-\dots$ by special Abelian methods. The series $1-1+1-\dots$ is summable (A) to $\frac{1}{2}$. It is instructive to consider its summability by other Abelian methods.

It is familiar in the theory of elliptic functions that

(4.10.1) $$1-2q+2q^4-2q^9+\dots = \prod\{(1-q^{2n+1})^2(1-q^{2n+2})\}$$

for $|q| < 1$, and the product plainly tends to 0 when $q \to 1$ by real values. It follows (writing e^{-x} for q) that

$$1-e^{-x}+e^{-4x}-e^{-9x}+\dots \to \tfrac{1}{2},$$

so that $1-1+1-\dots$ is summable (A, n^2) to $\frac{1}{2}$. It is also summable (A, n^k) for any positive k (compare Appendix V).

On the other hand, if $a > 1$, then

(4.10.2) $$F(x) = x-x^a+x^{a^2}-x^{a^3}+\dots$$

does not tend to a limit when $x \to 1$. To see this, we observe that $F(x)$ satisfies the functional equation

$$F(x)+F(x^a) = x,$$

and that $$\Phi(x) = \sum \frac{(-1)^n}{n!(1+a^n)}\left(\log\frac{1}{x}\right)^n$$

is another solution. Hence $\Psi(x) = F(x)-\Phi(x)$ satisfies $\Psi(x) = -\Psi(x^a)$, and is therefore a periodic function of $\log\log(1/x)$ with period $2\log a$. Since it is plainly not constant, it oscillates between finite limits of indetermination when $x \to 1$, $\log(1/x) \to 0$, $\log\log(1/x) \to -\infty$. But $\Phi(x) \to \frac{1}{2}$, and therefore $F(x)$ oscillates.

It follows that $1-1+1-\dots$ is not summable (A, λ) when $\lambda_n = a^n$ $(a > 1)$.

4.11. Lindelöf's and Mittag-Leffler's methods. There is one (A, λ) method which is particularly important in the theory of analytic continuation. In this

(4.11.1) $$\lambda_0 = 0, \qquad \lambda_n = n\log n \quad (n \geqslant 1).$$

If then $\sum a_n e^{-\lambda_n x} \to s$, we write $\sum a_n = s$ (L).

A power series $\sum a_n z^n$, convergent for small z, defines an analytic function $f(z)$ with a branch regular at the origin. In what follows we use $f(z)$ to denote this branch of the function, made uniform by an appropriate system of cuts in the plane of z. The 'Mittag-Leffler star' of $f(z)$ is the domain formed by drawing rays through 0 to every singular point of $f(z)$ and removing from the plane the parts of the rays beyond

the singular points. Thus the star of $\sum z^n = (1-z)^{-1}$ is the plane cut along the line $(1,\infty)$. The importance of the L method arises from the fact that it sums $\sum a_n z^n$ throughout the star of $f(z)$. We shall see later (§ 8.10) how this general theorem may be reduced to the special case in which $f(z)$ is the function $(1-z)^{-1}$: here we consider the special case only. It will be convenient to change our notation, writing δ for x.

THEOREM 32. *If D is any closed and bounded region which has no point on $(1,\infty)$, and λ_n is defined by* (4.11.1), *then*

$$\sum e^{-\delta\lambda_n} z^n \to (1-z)^{-1} \tag{4.11.2}$$

when $\delta \to 0$, uniformly in D.

We define $\Delta(\eta, R)$ as the region in the plane of $z = re^{i\theta}$ in which

$$0 < \eta \leqslant \theta \leqslant 2\pi-\eta, \qquad r \leqslant R; \tag{4.11.3}$$

and, since (4.11.2) is plainly true in any circle $r \leqslant 1-\zeta < 1$, it will be sufficient to show that

$$g_\delta(z) = \sum_1^\infty z^n e^{-\delta n \log n} \to z/(1-z) = g(z) \tag{4.11.4}$$

uniformly in Δ.

We define a contour C in the plane of $u = \rho e^{i\phi}$ by the circular arc $\rho = \frac{1}{2}$, $|\phi| \leqslant \phi_0 < \frac{1}{2}\pi$ and the two rays $\rho > \frac{1}{2}$, $|\phi| = \phi_0$. We shall suppose, as plainly we may, that ϕ_0 and δ_0 are chosen so that

$$\sin\phi_0 > \tfrac{1}{2}, \quad \tan\phi_0 > (4\log R)/\eta, \quad \delta_0\phi_0 < \tfrac{1}{2}\eta; \tag{4.11.5}$$

and we consider the integral

$$I_\delta(z) = \int z^u e^{-\delta u \log u} \frac{du}{e^{2\pi i u}-1}, \tag{4.11.6}$$

round C, where

$$z^u = e^{u \log z}, \qquad \log z = \log r + i\theta, \qquad \log u = \log\rho + i\phi,$$

and C is described so as to leave the origin on the right. Since

$$\begin{aligned} |e^{-\delta u \log u}| &= |\exp\{-\delta\rho(\cos\phi+i\sin\phi)(\log\rho+i\phi)\}| \\ &= \exp(-\delta\rho\log\rho\cos\phi+\delta\rho\phi\sin\phi), \end{aligned}$$

it follows from Cauchy's theorem that $I_\delta(z) = g_\delta(z)$ for $\delta > 0$ and z in Δ.† We now prove that $I_\delta(z)$ is uniformly convergent for $0 \leqslant \delta \leqslant \delta_0$ and z in Δ.

On the upper ray of C, we have

$$|e^{-\delta u \log u}| = \exp(-\delta\rho\log\rho\cos\phi_0+\delta\rho\phi_0\sin\phi_0),$$

† The integrand is dominated at infinity by the factor $\exp(-\delta\rho\log\rho\cos\phi)$.

$$|z^u| = |\exp\{\rho(\cos\phi_0+i\sin\phi_0)(\log r+i\theta)\}|$$
$$= \exp(\rho\log r\cos\phi_0-\rho\theta\sin\phi_0) \leqslant \exp(\rho\log R\cos\phi_0-\rho\eta\sin\phi_0),$$
$$\left|\frac{1}{e^{2\pi iu}-1}\right| \leqslant \frac{1}{1-e^{-2\pi\rho\sin\phi_0}} < \frac{1}{1-e^{-\frac{1}{2}\pi}}.$$

It follows from these inequalities and (4.11.5) that the integrand in (4.11.6) is majorized by a constant multiple of

$$\exp(\rho\log R\cos\phi_0-\tfrac{1}{2}\rho\eta\sin\phi_0) < e^{-\frac{1}{2}\rho\eta\sin\phi_0} < e^{-\frac{1}{4}\rho\eta},$$

and that this part of the integral is uniformly convergent.

The proof for the lower ray of C is similar.†

Since the integrand is uniformly continuous with respect to δ and z on any finite stretch of the contour, and $I_\delta(z)$ is regular in Δ for $\delta > 0$, it follows that

$$I_\delta(z) \to I(z) = \int\limits_C \frac{z^u}{e^{2\pi iu}-1}\,du$$

uniformly in Δ, and that the right-hand side is an analytic function of z regular in Δ. It is $g(z)$ when $-1 < z < 0$, and therefore throughout Δ; and the theorem follows.

There are other methods of summation, not (A, λ) methods, but of similar type to the L method, which have the same property. The most important is Mittag-Leffler's method, which we call M, and in which $\sum a_n$ is defined as

$$\lim_{\delta\to 0}\sum\frac{a_n}{\Gamma(1+\delta n)}.$$

It is easy to prove, by a variant of the method used above, with $\Gamma(1+\delta u)$ in the place of $e^{\delta u\log u}$, that $\sum z^n$ is summable (M) to $(1-z)^{-1}$ uniformly in Δ. The details naturally depend on the asymptotic properties of the gamma-function of a complex variable: an alternative proof will be given later.‡ Yet another method with similar properties is Le Roy's, in which $\sum a_n$ is defined as

$$\lim_{\zeta\to 1-0}\sum\frac{\Gamma(1+\zeta n)}{\Gamma(1+n)}a_n.$$

4.12. Means defined by integral functions. We consider next an important class of methods of which the best known is Borel's. Let us suppose that $J(x) = \sum p_n x^n$ is an integral function, not a polynomial, with non-negative coefficients p_n. If

$$\frac{S(x)}{J(x)} = \frac{\sum p_n s_n x^n}{\sum p_n x^n} \to s \tag{4.12.1}$$

† $|e^{2\pi iu}| = e^{2\pi\rho\sin\phi_0}$ is large for large ρ: $2\pi-\theta$ takes the place of θ in the argument, and $2\pi-\theta \geqslant \eta$.

‡ See p. 199 (note on § 8.10).

when $x \to \infty$, then we write

$$\sum a_n = s \text{ (J)}. \tag{4.12.2}$$

The simplest definition of this type is Borel's, in which $p_n = (n!)^{-1}$, $J(x) = e^x$: if

$$e^{-x}S(x) = e^{-x}\sum s_n \frac{x^n}{n!} \to s \tag{4.12.3}$$

when $x \to \infty$, then we write

$$\sum a_n = s \text{ (B)}. \tag{4.12.4}$$

There is an alternative definition of Borel which we shall consider in § 4.13 and Ch. VIII.

Here
$$t(x) = \sum c_n(x)s_n,$$
$$c_n(x) = p_n x^n / \sum p_n x^n \geqslant 0, \quad c_n(x) \to 0, \quad \sum |c_n(x)| = \sum c_n(x) = 1,$$
and the conditions of Theorem 5 are plainly satisfied. Thus

THEOREM 33. *The* J *method is regular.*

The J method provides a convenient opportunity for a more explicit statement of a general principle which we have referred to already† and of which we shall find many applications later. A method may be said to be 'powerful' if it can sum rapidly divergent series: thus Borel's method is more powerful than the (C, 1) or A methods, which will not sum $\sum z^n$ outside its circle of convergence. Borel's method, on the other hand, sums it in the half-plane $\Re z < 1$. For in this case $s_n = (1-z^{n+1})/(1-z)$, and

$$\frac{S(x)}{J(x)} = e^{-x}\sum \frac{1-z^{n+1}}{1-z}\frac{x^n}{n!} = \frac{1}{1-z} - \frac{ze^{-(1-z)x}}{1-z} \to \frac{1}{1-z},$$

provided only that $\Re z < 1$. In particular it sums the series for all negative z.

In this sense the J method is the more powerful the more rapidly p_n tends to 0. Thus Borel's method sums $1-a+a^2-a^3+...$ for all positive a, but it will not sum the series for which $s_n = (-1)^n n!\, a^n$ because $\sum (-1)^n (ax)^n$ is not convergent when $ax > 1$. If we take this s_n, and $p_n = (n!)^{-2}$, then

$$S(x) = \sum (-1)^n \frac{(ax)^n}{n!} = e^{-ax}, \qquad J(x) = \sum \frac{x^n}{(n!)^2} = I_0(2\sqrt{x}),‡$$

and $S(x)/J(x) \to 0$.

We shall, however, find that, usually, *the delicacy of a method decreases as its power increases*, and that very powerful methods, adapted to the

† See § 3.8.

‡ $I_0(x)$ being Bessel's function with imaginary argument.

summation of rapidly divergent series, are apt to fail with divergent series of a less violent kind (such as we encounter, for example, in the theory of Fourier series). Thus we shall find in §4.15 that the J method with $p_n = e^{-cn^2}$ fails to sum $1-1+1-\dots$.

We supposed in (4.12.1) that $J(x)$ is an integral function and that $x \to \infty$. There is a modification in which the radius of convergence of (4.12.1) is finite. In this case we may take it to be 1, and must suppose $\sum p_n$ divergent. The definition is still expressed by (4.12.1), but now $x \to 1$, and the method then resembles the 'Abelian' methods of §§ 4.7–10. Thus, if $p_n = 1$, then $J(x) = (1-x)^{-1}$, and the definition becomes $(1-x)\sum s_n x^n \to s$, i.e. $\sum a_n x^n \to s$. This is the A definition.

If $p_n = (n+1)^{-1}$, the definition becomes

$$\left(\log\frac{1}{1-x}\right)^{-1}\sum \frac{s_n}{n+1}x^{n+1} \to s$$

or

$$\int_0^x \frac{f(t)}{1-t}\,dt \sim s\log\frac{1}{1-x}, \tag{4.12.5}$$

where $f(x) = \sum a_n x^n$. It is plain that $f(x) \to s$ implies (4.12.5), so that the method includes the A method.

4.13. Moment constant methods. A moment constant μ_n is a number of the form

$$\mu_n = \int_0^\infty x^n\,d\chi, \tag{4.13.1}$$

where $\chi = \chi(x)$ is a bounded and increasing function of x such that the Stieltjes integral (4.13.1) converges for all n. If ξ is the lower bound of numbers x for which

$$\int_x^\infty d\chi(u) = 0,$$

then

$$\int_{\xi+0}^\infty d\chi(u) = 0, \qquad \int_x^\infty d\chi(u) > 0 \quad (x < \xi)$$

and

$$\mu_n = \int_0^\infty x^n\,d\chi = \int_0^{\xi-0} x^n\,d\chi + \{\chi(\xi+0)-\chi(\xi-0)\}\xi^n,$$

when ξ is finite. Usually, however, ξ will be ∞; and we shall suppose, when ξ is finite, that χ is continuous at ξ.†

If

$$a(x) = \sum (a_n/\mu_n)x^n, \tag{4.13.2}$$

† See the note at the end of the chapter.

then formal term-by-term integration gives

$$\int a(x)\,d\chi = \sum (a_n/\mu_n) \int x^n\,d\chi = \sum a_n,$$

and this suggests that we take the integral on the left as the basis of a definition of the sum of $\sum a_n$.

We write

$$\int a(x)\,d\chi = s \tag{4.13.3}$$

if either (i) $\xi = \infty$, the series (4.13.2) converges for all x, and

$$\lim_{X\to\infty} \int_0^X a(x)\,d\chi = s,$$

or (ii) $\xi < \infty$, $\chi(\xi+0)-\chi(\xi-0) = 0$, (4.13.2) converges for $0 < x < \xi$, and

$$\int_0^{\xi-0} a(x)\,d\chi = \lim_{X\to\xi-0} \int_0^X a(x)\,d\chi = s;$$

and

$$\sum a_n = s \quad (\mu_n) \tag{4.13.4}$$

in either of these two cases.

THEOREM 34. *The* (μ_n) *method is regular.*

If $\sum a_n$ is convergent and $0 < X < X_1 < \xi \leqslant \infty$, then

$$\mu_n = \int_0^\infty x^n\,d\chi \geqslant X_1^n \int_{X_1}^\infty d\chi > 0 \tag{4.13.5}$$

and (4.13.2) converges uniformly for $0 \leqslant x \leqslant X$. Also

$$\left|\frac{s_n}{\mu_n}\int_0^X x^n\,d\chi\right| \leqslant |s_n|\left(\frac{X}{X_1}\right)^n \int_0^\infty d\chi \Big/ \int_{X_1}^\infty d\chi \to 0 \tag{4.13.6}$$

when $n\to\infty$, for $X < X_1$ and so for $X < \xi$. In particular, taking $s_n = 1$,

$$\frac{1}{\mu_n}\int_0^X x^n\,d\chi \to 0. \tag{4.13.7}$$

It follows from the uniform convergence of (4.13.2) that

$$t(X) = \int_0^X \left(\sum \frac{a_n}{\mu_n} x^n\right) d\chi = \sum \frac{a_n}{\mu_n} \int_0^X x^n\,d\chi,$$

and from (4.13.6) that

$$t(X) = \sum s_n\left(\frac{1}{\mu_n}\int_0^X x^n\,d\chi - \frac{1}{\mu_{n+1}}\int_0^X x^{n+1}\,d\chi\right) = \sum c_n(X)s_n, \tag{4.13.8}$$

say. Plainly $c_n(X) \to 0$ when $X \to \xi$. Also

$$\int_0^\infty x^{n+1}\,d\chi \int_0^X x^n\,d\chi - \int_0^X x^{n+1}\,d\chi \int_0^\infty x^n\,d\chi$$

$$= \int_X^\infty x^{n+1}\,d\chi \int_0^X x^n\,d\chi - \int_0^X x^{n+1}\,d\chi \int_X^\infty x^n\,d\chi$$

$$\geqslant X\left(\int_X^\infty x^n\,d\chi \int_0^X x^n\,d\chi - \int_0^X x^n\,d\chi \int_X^\infty x^n\,d\chi\right) = 0.\dagger$$

Hence $$c_n(X) \geqslant 0,$$

$$\sum |c_n(X)| = \sum c_n(X) = \frac{1}{\mu_0}\int_0^X d\chi - \lim_{n\to\infty}\frac{1}{\mu_n}\int_0^X x^n\,d\chi = \frac{1}{\mu_0}\int_0^X d\chi,$$

by (4.13.7), and $\sum c_n(X) \to 1$ when $X \to \xi$. Thus the conditions of Theorem 5 are satisfied, and the method is regular.

The most important case is that in which

$$\chi(x) = 1 - e^{-x^{1/\alpha}} \quad (\alpha > 0).$$

Then $$\mu_n = \frac{1}{\alpha}\int e^{-x^{1/\alpha}} x^{(1/\alpha)-1} x^n\,dx = \int e^{-u} u^{n\alpha}\,du = \Gamma(n\alpha+1),$$

and the definition is

$$\int e^{-u} \sum \frac{a_n u^{n\alpha}}{\Gamma(n\alpha+1)}\,du = s. \tag{4.13.9}$$

In these circumstances we write

$$\sum a_n \doteq s \ (\mathrm{B}', \alpha). \tag{4.13.10}$$

In particular, when $\alpha = 1$, we write

$$\sum a_n = s \ (\mathrm{B}'). \tag{4.13.11}$$

We shall see in Ch. VIII that the definitions (4.13.11) and (4.12.4) are intimately connected and 'all but' equivalent. We were led to them in quite different ways, and their close connexion is due to the special properties of the exponential function.

If $\alpha = 1$, $a_n = z^n$, then

$$\int e^{-x} \sum \frac{(zx)^n}{n!}\,dx = \int e^{-x(1-z)}\,dx = \frac{1}{1-z}$$

when $\Re z < 1$. Thus the B′ method, like the B method, sums $\sum z^n$ in this half-plane.

† If $\xi < \infty$, the upper limit ∞ may be replaced by $\xi - 0$ throughout.

We make the moment method more 'powerful' by increasing the magnitude of μ_n. Such an increase of power carries its disadvantages with it. Thus, if

(4.13.12) $$d\chi = e^{-k(\log x)^2}x^{-1}\,dx \quad (k > 0),$$

then

(4.13.13) $$\mu_n = \int\limits_0^\infty e^{-k(\log x)^2}x^{n-1}\,dx = \int\limits_{-\infty}^{\infty} e^{-ku^2+nu}\,du = \sqrt{\left(\frac{\pi}{k}\right)}e^{n^2/4k},$$

and the definition becomes

(4.13.14) $$\sqrt{\left(\frac{k}{\pi}\right)} \int\limits_{-\infty}^{\infty} e^{-ku^2}\left(\sum e^{nu-n^2/4k}a_n\right)du = s.$$

We shall see in § 4.15 that this method will not sum $1-1+1-\dots$.

If $\xi = 1$, $\chi = x$ for $x < 1$, $\chi = 1$ for $x \geqslant 1$, then $\mu_n = (n+1)^{-1}$ and the definition is

$$\lim_{\epsilon \to 0} \int\limits_0^{1-\epsilon} \{\sum (n+1)a_n x^n\}\,dx = s.$$

This is plainly equivalent to the A definition.

4.14. A theorem of consistency. There is no general theorem of consistency for moment constant methods: different methods may sum the same series to different sums. But there is a special theorem of consistency which is sometimes useful, in which we suppose that $\xi = \infty$ and

(4.14.1) $$\chi(x) = \int\limits_0^x \phi(t)\,dt.$$

THEOREM 35. *Suppose* (i) *that $\phi(x)$ is positive and decreasing;* (ii) *that*

$$\mu_n = \int x^n \phi(x)\,dx$$

is convergent for $n \geqslant 0$; and (iii) *that $\phi(\zeta x)/\phi(x)$ is, for every fixed $\zeta > 1$, a decreasing function of x; or at any rate that conditions* (i) *and* (iii) *are satisfied for $x > x_0$. Suppose further* (iv) *that $\sum a_n z^n$ is convergent for small z; and* (v) *that*

(4.14.2) $$\int \left(\sum \frac{a_n}{\mu_n} x^n\right)\phi(x)\,dx = s,$$

so that $\sum a_n$ is summable (μ_n) to s. Then $\sum a_n z^n$ is uniformly summable for $0 \leqslant z \leqslant 1$; so that it represents an analytic function $f(z)$, which is regular on the segment $(0, 1)$ *and tends to s when $z \to 1$ through real values less than* 1.

This is a theorem of consistency because it shows that the sum s is fixed by the function $f(z)$ independently of the special $\phi(x)$ and μ_n used in the definition.†

It is plainly sufficient to prove the series uniformly summable in any interval $0 < \delta \leqslant z \leqslant 1$. The series $g(x) = \sum (a_n/\mu_n)x^n$ converges for all x, and

$$\int g(x)\phi(x)\,dx = s,$$

by (4.14.2); and the sum of $\sum a_n z^n$ is

(4.14.3) $$\int g(zx)\phi(x)\,dx = \frac{1}{z}\int g(x)\phi\left(\frac{x}{z}\right)dx,$$

† Compare the second proof of Theorem 17 in § 4.2.

if this integral is convergent. Now if $X > x_0$ then

$$\int_X^{X'} g(x)\phi\left(\frac{x}{z}\right)dx = \int_X^{X'} g(x)\phi(x)\frac{\phi(x/z)}{\phi(x)}\,dx = \frac{\phi(X/z)}{\phi(X)}\int_X^{X''} g(x)\phi(x)\,dx,$$

where $X < X'' < X'$, by condition (iii). The outside factor does not exceed 1, and the second is numerically less than ϵ for $X \geqslant X_0(\epsilon)$. Hence the integral (4.14.3) converges uniformly for $0 < \delta \leqslant z \leqslant 1$, and the theorem follows.

The conditions are satisfied, for example, if $\phi(x) = e^{-Ax^\alpha}x^{a-1}$, where $A > 0$, $\alpha > 0$, $a > 0$.

As an application, suppose that

$$\sum a_n z^n = 1 - az + \frac{a(a+1)}{2!}z^2 - \dots = (1+z)^{-a}$$

for small z. If we take $\phi(x) = e^{-x}$, as in Borel's method, we obtain the sum

$$\int e^{-x}\left\{1 - \frac{a}{(1!)^2}xz + \frac{a(a+1)}{(2!)^2}x^2z^2 - \dots\right\}dx,$$

and the value of the integral is not obvious. It is much more convenient to take $\phi(x) = x^{a-1}e^{-x}$, when $\mu_n = \Gamma(n+a)$ and we obtain

$$\frac{1}{\Gamma(a)}\int e^{-x}x^{a-1}\left(1 - \frac{xz}{1!} + \frac{x^2z^2}{2!} - \dots\right)dx = \frac{1}{\Gamma(a)}\int e^{-x(1+z)}x^{a-1}\,dx = (1+z)^{-a},$$

provided only that $\Re z > -1$.

4.15. Methods ineffective for the series $1-1+1-\dots$. In this section we illustrate the general principle stated in § 4.12 by showing how two 'violent' methods, one of 'integral function' and one of 'moment constant' type, fail to sum $1-1+1-\dots$.†

(1) Let us take $p_n = e^{-cn^2}$, where $c > 0$, in the definition (4.12.1), and write e^u for x. Then, since $s_{2m} = 1$ and $s_{2m+1} = 0$, we have to determine whether

$$\sum e^{-4cm^2+2um} \Big/ \sum e^{-cn^2+un}$$

tends to a limit when $u \to \infty$. It is plain that we may replace this ratio by $F_1(u)/F_2(u)$, where F_1 and F_2 are the sums extended from $-\infty$ to ∞. Now

$$F_1(u) = \vartheta_3\left(\frac{iu}{\pi}\middle|\frac{4ic}{\pi}\right), \qquad F_2(u) = \vartheta_3\left(\frac{iu}{2\pi}\middle|\frac{ic}{\pi}\right),$$

and

$$\vartheta_3(v+n\tau|\tau) = e^{-n^2\pi i\tau - 2n\pi iv}\vartheta_3(v|\tau);$$

and it follows from these formulae that $F_1(u)/F_2(u)$ has the period $4c$. Since it is plainly not constant, it does not tend to a limit.

(2) Let us suppose χ and μ_n defined as in (4.13.12) and (4.13.13). Then the sum of $1-1+1-\dots$ is defined as

$$\sqrt{\left(\frac{k}{\pi}\right)}\int_{-\infty}^{\infty} e^{-ku^2}\left(\sum_{n=0}^{\infty}(-1)^n e^{-(n^2/4k)+nu}\right)du, \tag{4.15.1}$$

if this integral is convergent; and it is plain that the convergence will not be affected by replacing the lower limit of summation by $-\infty$. But

$$F(u) = \sum_{n=-\infty}^{\infty}(-1)^n e^{-(n^2/4k)+nu} = \vartheta_4\left(\frac{iu}{2\pi}\middle|\frac{i}{4k\pi}\right) = \vartheta_4(v|\tau),$$

† See also § 4.10, for the failure of a violent method of Abelian type.

and it is easily verified that

$$\int_{n/2k}^{(n/2k)+a} e^{-ku^2}F(u)\,du = (-1)^n \int_0^a e^{-ku^2}F(u)\,du$$

for any a. It follows that the integral (4.15.1) is not convergent and the series not summable.

4.16. Riesz's typical means. The 'typical means' of M. Riesz are generalizations of certain means which we shall consider in § 5.16, and a full discussion of them would be more in place in a book dealing specially with the theory of Dirichlet's series. We therefore dismiss them very shortly here.

Suppose that λ_n satisfies (4.7.1), that

$$A_\lambda(x) = a_0+a_1+\dots+a_n = s_n \quad (\lambda_n < x \leqslant \lambda_{n+1}), \qquad A_\lambda(x) = 0 \quad (x \leqslant \lambda_0),$$

that $\kappa > 0$ and that

$$A_\lambda^{(\kappa)}(\omega) = \frac{\kappa}{\omega^\kappa}\int_0^\omega A_\lambda(x)(\omega-x)^{\kappa-1}\,dx \to s \tag{4.16.1}$$

when $\omega \to \infty$. Then we say that $\sum a_n$ is summable (R, λ, κ) to s. We have also, by partial integration,

$$A_\lambda^{(\kappa)}(\omega) = \frac{1}{\omega^\kappa}\int_0^\omega (\omega-x)^\kappa\,dA_\lambda(x) = \sum_{\lambda_n \leqslant \omega}\left(1-\frac{\lambda_n}{\omega}\right)^\kappa a_n.$$

We can write (4.16.1) as

$$A_\lambda^{(\kappa)}(\omega) = \int \phi(x,\omega)A_\lambda(x)\,dx,$$

where

$$\phi = \kappa\omega^{-\kappa}(\omega-x)^{\kappa-1} \quad (0 \leqslant x < \omega), \qquad \phi = 0 \quad (x \geqslant \omega).$$

Then $\phi \geqslant 0$, $\phi = O(\omega^{-1})$ for large ω, uniformly in any finite interval of x, and

$$\int_0^\infty |\phi(x,\omega)|\,dx = \frac{\kappa}{\omega^\kappa}\int_0^\omega (\omega-x)^{\kappa-1}\,dx = 1.$$

Hence, after Theorem 6,

THEOREM 36. *Riesz's typical means are regular.*

It is easily verified that the $(R, n, 1)$ method is equivalent to the $(C, 1)$ method. We shall prove more than this in § 5.16.

Another interesting case is that in which $\lambda_n = \log(n+1)$, $\kappa = 1$. We prove

THEOREM 37. *In order that* $\sum a_n$ *should be summable* $(\mathrm{R}, \lambda, 1)$, *with* $\lambda_n = \log(n+1)$, *to* s, *it is necessary and sufficient that*

$$\frac{1}{\log(n+1)}\left(s_0+\frac{s_1}{2}+\frac{s_2}{3}+\ldots+\frac{s_n}{n+1}\right) \to s. \tag{4.16.2}$$

In other words, the means are then equivalent to the logarithmic means of § 3.8.

If $\omega = \log(q+1)$, $n = [q]$, then the definition (4.16.1) reduces to

$$\frac{1}{\log(q+1)} \sum_0^{n-1} s_m \log\frac{m+2}{m+1} + \frac{s_n}{\log(q+1)} \log\frac{q+1}{n+1} \to s, \tag{4.16.3}$$

and we have to prove this equivalent to (4.16.2).

Let us assume (4.16.2), and write

$$u_n = \frac{s_n}{n+1}, \qquad U_n = u_0+u_1+\ldots+u_n.$$

Then $U_n \sim s\log n$, so that $u_n = o(\log n)$ and $s_n = o(n\log n)$. Hence the last term in (4.16.3) tends to zero. Also

$$\log\frac{m+2}{m+1} = \frac{1}{m+1} - \frac{1}{2(m+1)^2} + O\left(\frac{1}{m^3}\right),$$

so that the sum in (4.16.3) is

$$\sum_0^{n-1} u_m - \frac{1}{2}\sum_0^{n-1}\frac{u_m}{m+1} + \sum_0^{n-1} u_m\, O\left(\frac{1}{m^2}\right) = P_n - \tfrac{1}{2}Q_n + R_n,$$

say. Here

$$Q_n = \sum_0^{n-1}\frac{u_m}{m+1} = \sum_0^{n-2}\frac{U_m}{(m+1)(m+2)} + \frac{U_{n-1}}{n} = O(1)$$

because $U_n = O(\log n)$, and R_n is plainly $O(1)$. Hence (4.16.3) reduces to $P_n \sim s\log(q+1)$, which is equivalent to (4.16.2).

The proof of the converse is similar but simpler, since we may take $\omega = \log(n+1)$.

4.17. Methods suggested by the theory of Fourier series. The series

$$\tfrac{1}{2}+\cos\theta+\cos 2\theta+\ldots = \sum{}' \cos n\theta\dagger \tag{4.17.1}$$

is fundamental in the theory of Fourier series. Its partial sum is

$$s_n(\theta) = \sum_0^{n}{}' \cos\nu\theta = \frac{\sin(n+\frac{1}{2})\theta}{2\sin\frac{1}{2}\theta} = D_n(\theta), \tag{4.17.2}$$

† The dash implying that the term with $n = 0$ has a factor $\frac{1}{2}$.

and its means defined by (3.1.3) and (3.1.4) are

(4.17.3) $t_m(\theta) = \sum c_{m,n} D_n(\theta)$, (4.17.4) $t(x,\theta) = \sum c_n(x) D_n(\theta)$.

In particular the (C, 1) and A means are

$$t_m(\theta) = \frac{1}{2(m+1)}\left\{\frac{\sin\frac{1}{2}(m+1)\theta}{\sin\frac{1}{2}\theta}\right\}^2, \tag{4.17.5}$$

$$t(r,\theta) = \frac{1-r^2}{2(1-2r\cos\theta+r^2)}.\dagger \tag{4.17.6}$$

The $t_m(\theta)$ defined by (4.17.5) has the properties that

$$t_m(\theta) \geqslant 0, \qquad \frac{1}{\pi}\int_{-\pi}^{\pi} t_m(\theta)\,d\theta = 1,$$

and that $t_m(\theta) \to 0$, when $m \to \infty$, uniformly in any closed sub-interval of $(-\pi, \pi)$ which does not include the origin; and the $t(r,\theta)$ of (4.17.6) has similar properties. It is on these properties that the applications of the methods to Fourier series are based, and other choices of a $t_m(\theta)$ with the same properties lead to valuable methods of summation. Thus

$$t_m(\theta) = \frac{2^{-m-1}\Gamma(m+1)\sqrt{\pi}}{\Gamma(m+\frac{1}{2})}(1+\cos\theta)^m$$

has the properties required. Since

$$(1+\cos\theta)^m = 2^{1-m}\frac{2m!}{(m!)^2}\left\{\tfrac{1}{2}+\frac{m}{m+1}\cos\theta+\frac{m(m-1)}{(m+1)(m+2)}\cos 2\theta+\dots\right\},$$

we are led to de la Vallée-Poussin's definition (VP)

$$\sum a_n = \lim_{m\to\infty}\left\{a_0+\frac{m}{m+1}a_1+\frac{m(m-1)}{(m+1)(m+2)}a_2+\dots\right\}.$$

In terms of s_n

$$t_m = \frac{1}{m+1}s_0+\frac{3m}{(m+1)(m+2)}s_1+\frac{5m(m-1)}{(m+1)(m+2)(m+3)}s_2+\dots,$$

and it is easily verified that the method is regular.

In these methods the coefficient of s_n is non-negative. There are other methods, important in the theory of general trigonometrical series, in which this is not so. The most fundamental is Riemann's, in which we define $\sum a_n$ as

$$\lim_{h\to 0} t(h) = \lim_{h\to 0}\sum a_n\left(\frac{\sin nh}{nh}\right)^2:$$

† It is usual to write r for x here, and h for x in the 'Riemann' definitions.

the coefficient of a_0 is interpreted as 1. It is familiar that this method, usually called the (R, 2) method, is regular. In this case

$$t(h,\theta) = \frac{\pi(2h-|\theta|)}{4h^2} \quad (|\theta| \leqslant 2h), \qquad 0 \quad (2h \leqslant \theta \leqslant \pi)$$

and $t(h,\theta)$ has the properties corresponding to those of (4.17.5) and (4.17.6).

More generally, summability (R, k), where k is a positive integer, is defined by

$$\sum a_n\left(\frac{\sin nh}{nh}\right)^k \to s.$$

The method is regular for $k > 1$ but not for $k = 1$.

Another method, closely connected with the (R, 2) method, but not equivalent to it, is the (R_2) method defined by

$$t(h) = \frac{2}{\pi}\sum \frac{\sin^2 nh}{n^2 h}\, s_n,$$

where the coefficient of s_0 in the sum is interpreted as h. This method also is regular.

4.18. A general principle. Most of the definitions which we have considered in the preceding sections may be presented as illustrations of a general principle.

Let us suppose that $F = F(\alpha, \beta, \gamma, \ldots)$ is a function of certain parameters α, β, $\gamma, \ldots$ which tend to limits α_0, β_0, $\gamma_0, \ldots$; that A, B, $C, \ldots$ are the limit operations $\alpha \to \alpha_0$, $\beta \to \beta_0$, $\gamma \to \gamma_0, \ldots$; that

$$PF = ABC\ldots F = \lim_{\alpha\to\alpha_0}\left\{\lim_{\beta\to\beta_0}\left(\lim_{\gamma\to\gamma_0}\ldots\right)\right\}F;$$

and $QF = A'B'C'\ldots F$, where A', B', $C', \ldots$ are A, B, $C, \ldots$ in a different order. We may ask whether

$$PF = QF, \tag{4.18.1}$$

and the theorems which assert that this is true under appropriate conditions include many of the most important in analysis.

We may also look at the equation (4.18.1) from a different point of view. Suppose, for example, that

$$F = F(n,x) = \sum_0^n a_m x^m,$$

that $\alpha = n$, $\beta = x$, and that A and B are the operations $n \to \infty$, $x \to 1$. Then

$$BF = \lim_{x\to1}\sum_0^n a_m x^m = \sum_0^n a_m,$$

and $$PF = ABF = \lim_{n\to\infty} \sum_0^n a_m = \sum_0^\infty a_m$$

if and only if $\sum a_m$ is convergent. On the other hand,

$$AF = \lim_{n\to\infty} \sum_0^n a_m x^m = \sum_0^\infty a_n x^n = f(x),$$

say, whenever this last series is convergent for $x < 1$; and

$$QF = BAF = \lim f(x)$$

is Abel's limit for the series $\sum a_m$. If PF exists, then QF exists and is equal to PF, but QF exists in many cases in which PF does not. In these circumstances we may take QF as the *definition* of the symbol PF, and agree to write PF when we mean QF. The utility of such a fiction is, of course, to be judged by its results.

Again, for the J definition of § 4.12, with $p_n > 0$ for all n,

$$F = \Big(\sum_0^n p_m s_m x^m\Big)\Big/\Big(\sum_0^n p_m x^m\Big),$$

A is $n \to \infty$, B is $x \to \infty$; $BF = s_n$, $ABF = s$ if and only if $s_n \to s$; and BAF is the limit which we took as our definition. For the 'moment constant' definition of § 4.13 (with $\xi = \infty$),

$$F = \int_0^X \Big(\sum_0^n \frac{a_m}{\mu_m} x^m\Big)\, d\chi = \sum_0^n \frac{a_m}{\mu_m} \int_0^X x^m\, d\chi;$$

A is $n \to \infty$, B is $X \to \infty$;

$$BF = \sum_0^n \frac{a_m}{\mu_m} \int_0^\infty x^m\, d\chi = \sum_0^n a_m,$$

so that $ABF = s$ if and only if $\sum a_n$ converges to s; and

$$BAF = \lim_{X\to\infty} \int_0^X a(x)\, d\chi = \int_0^\infty a(x)\, d\chi.$$

We may sometimes wish to connect the operations $A, B, \ldots$ by relations between $\alpha, \beta, \ldots$. Suppose, for example, that

$$F = F(n,p) = \sum_0^n \Big(1 - \frac{m}{n+1}\Big)^{(n+1)/p} a_m.$$

Then $$\lim_{p\to\infty} F = \sum_0^n a_m, \qquad \lim_{n\to\infty} \lim_{p\to\infty} F = \sum_0^\infty a_m$$

when the series is convergent; and it is easy to prove that, when $n \to \infty$, $F \to \sum e^{-m/p} a_m$ whenever the last series is convergent; so that

$$\lim_{p\to\infty} \lim_{n\to\infty} F = \lim_{p\to\infty} \sum e^{-m/p} a_m,$$

which is Abel's limit, whenever this limit exists. Thus the ordinary sum and Abel's limit correspond to the two repeated limits of F.

On the other hand, we may make n and p tend to infinity together. For example, if we suppose that $n+1 = p$, then

$$F = \sum_0^n \left(1-\frac{m}{n+1}\right) a_m = \frac{s_0+s_1+\ldots+s_n}{n+1},$$

and we obtain the (C, 1) definition. If we suppose that $n+1 = kp$, we obtain what is very nearly the (C, k) definition of Ch. V.†

NOTES ON CHAPTER IV

§ 4.1. The general definition of a 'Nörlund mean' occurs first in Voronoi, *Proc. of the eleventh congress of Russian naturalists and scientists* (in Russian), St. Petersburg, 1902, 60–1. There is an annotated English translation by Tamarkin, *Annals* (2), 33 (1932), 422–8. Voronoi's article was a short note in a rare publication, and was unnoticed until Tamarkin called attention to it. A number of special cases of the definition, such as Cesàro's, were, of course, already familiar.

Nörlund gave the definition independently in *Lunds Universitets Årsskrift* (2), 16 (1920), no. 3. He (explicitly) and Voronoi (tacitly) assume that $p_n/P_n \to 0$, so that the method is regular.

§ 4.2. Of the two proofs of Theorem 17, the first is Nörlund's. The second, depending on Theorem 18, was given independently by Zygmund, *Mathesis Polska*, 1 (1926), 75–85 and 119–29, and by Silverman and Tamarkin, *MZ*, 29 (1928), 161–70. Voronoi states the theorem, and his short indications show that his proof was on the lines followed by these later authors.

§§ 4.3–4. Theorems 19 and 21 are due to M. Riesz, *PLMS* (2), 22 (1923), 412–19.

The condition (4.3.7) is also unnecessary when both $\sum p_n$ and $\sum q_n$ are convergent, but the question remains open when $\sum p_n < \infty$, $\sum q_n = \infty$.

§ 4.5. Theorem 22 is proved, with a different purpose, by Szegö, *MZ*, 25 (1926), 172–87 (177). Szegö attributes the result to Kaluza. Theorem 23 seems to be new. I had originally inserted the additional condition $p_n = o(q_n)$, but Dr. Bosanquet showed me that this condition is unnecessary.

§§ 4.7–8. Theorem 26, and a generalization for complex b_n, were proved by Jensen, *CR*, 103 (1886), 980 and 106 (1888), 835. See Pringsheim, *Vorlesungen über Zahlen- und Funktionentheorie* (Leipzig, 1916), 1, 308–10 and 938.

Theorems of the type of Theorems 25 and 29 are familiar, and have been generalized by many writers in many directions. For these particular theorems see Dienes, 394–7; Hardy, *PLMS* (2), 4 (1906), 247–65; and Perron, *MZ*, 6 (1920), 286–310. We can prove, a little more generally, that (4.7.4) and (4.7.6) are necessary and sufficient conditions for $\sum a_n = s$ to imply $\phi(x) \to s$.

† See in particular § 5.16.

Theorems 28 and 30 were proved by Hardy in *MM*, 39 (1910), 136–9 and *PLMS* (2), 8 (1910), 301–20 (318).

§ 4.9. For $\lambda_n = n$, Stolz, *Zeitschrift für Math.* 29 (1884), 127–8: see Stolz und Gmeiner, *Einleitung in die Funktionentheorie* (Leipzig, 1905), 2, 287–8, or Bromwich, 252–5. For general λ_n, Cahen, *AEN* (3), 11 (1894), 75–164 (86–7): see Landau, *Handbuch*, 737–8, or Hardy and Riesz, 3–4.

§ 4.10. For (4.10.1) see Tannery and Molk, 2, 10–13, or Hardy and Wright, 280–2.

Hardy, *QJM*, 38 (1907), 269–88, discusses the series (4.10.2) in detail, and proves the formula

$$e^{-y}-e^{-ay}+e^{-a^2y}-\dots = \sum_0^\infty \frac{(-1)^n}{a^n+1}\frac{y^n}{n!}+\frac{1}{\log a}\sum_{-\infty}^{\infty}\Gamma\left\{-\frac{(2n+1)\pi i}{\log a}\right\}y^{(2n+1)\pi i/\log a},$$

which shows the oscillations when $y \to 0$ explicitly. The argument here is due to Maclagan-Wedderburn.

§ 4.11. The appropriate references to the work of Le Roy, Lindelöf, and Mittag-Leffler are given in the note on § 8.10. The proof of Theorem 32 is Lindelöf's.

§ 4.12. Borel gave the general definition (4.12.1) in his earliest work on the subject: see Borel, 95. The regularity of the B definition was proved first by Hardy, *TCPS*, 19 (1902), 297–321 (298–300).

§ 4.13. There is a very clear account of the simpler properties of the Stieltjes integral in Widder, ch. 1.

Theorem 34 is proved by Good, *JLMS*, 19 (1944), 141–3, except that he supposes χ absolutely continuous. We have ignored the case

$$\xi < \infty, \quad \chi(\xi+0)-\chi(\xi-0) = D > 0,$$

which actually leads to a 'trivial' method, i.e. one summing convergent series only. In *JLMS*, 21 (1946), 110–18, Good proves a further theorem of the same character.

§ 4.14. Theorem 35 is a corrected version of one stated by Bromwich (1), 301–2. The conditions which he gives are unnecessarily strong in one way and inadequate in another. The example at the end of the section is his.

Mr. Eggleston observes that we can dispense with condition (iii) if the integral in (v) is absolutely convergent.

§ 4.15. The formulae used for the transformation of theta-functions will be found in Tannery and Molk, 2, 263 (Table XLIII).

§ 4.16. For the general theory of Riesz's typical means see Hardy and Riesz.

§ 4.17. There are general accounts of the theory of the summation of Fourier series in Hardy and Rogosinski, ch. 5, and Zygmund, ch. 3.

De la Vallée-Poussin's method (VP) was defined by him in *Bulletin de l'Acad. Sc. de Belgique* (1908), 193–254, and applied to the summation of the successive derived series of Fourier series. Gronwall, *JM*, 147 (1917), 16–35, proved that any series summable (C, k) is summable (VP). He also proved that the series $\sum z^n$ is summable (VP) to $1/(1-z)$ in the interior of the outer loop of the limaçon

(1) $$|1+z|^2 = 4|z|,$$

from which it follows that the VP method is stronger than the aggregate of the (C, k) methods.

The VP method has very close relations to the $(A, 2)$ method. Thus Hardy

(l.c. under § 2.8) proved that the methods are equivalent for Fourier series; and Hyslop, *PLMS* (2), 40 (1936), 449–67, extended the equivalence to all series for which $a_n = O(n^K)$. He also observed that $\sum z^n$ is summable (A, 2) inside the curve

$$r = e^{|\theta|} \quad (|\theta| \leqslant \pi), \tag{2}$$

which includes the curve (1), except for the point $z = 1$, in its interior, so that there are series summable (A, 2) but not summable (VP). Later, Kuttner, *PLMS* (2), 44 (1938), 92–9, proved that (VP) $\rightarrow$ (A, 2) in all cases.

Another method with very similar properties has been defined by Obrechkoff, *CR*, 182 (1926), 307–9.

The Riemann methods are fundamental in the theory of trigonometrical series. Thus the regularity of (R, 2) is 'Riemann's first lemma' and that of (R_2) is his second. A good deal has been written recently about the relations of (R, k) and (C, l). Thus Verblunsky, *PCPS*, 26 (1930), 34–42, proved the implication

$$(\mathrm{C}, k-\delta) \rightarrow (R, k+1),$$

and Kuttner, *PLMS* (2), 38 (1935), 273–83, proved $(\mathrm{R}, 1) \rightarrow (\mathrm{C}, 1+\delta)$ and $(\mathrm{R}, 2) \rightarrow (\mathrm{C}, 2+\delta)$: here δ is any positive number. Kuttner gives other references.

Marcinkiewicz, *JLMS*, 10 (1935), 268–72, proves the 'incomparability' of (R, 2) and (R_2). See also Kuttner, *PLMS* (2), 40 (1936), 524–40; Hardy and Rogosinski, *PCPS*, 43 (1947), 10–25 (where it is shown that the methods are incomparable even for Fourier series).

§ 4.18. For all this see Hardy and Chapman, *QJM*, 42 (1911), 181–215.

V

ARITHMETIC MEANS (1)

5.1. Introduction. The simplest method of summation of a divergent series is the first method of § 1.3. There are many important generalizations of this method, and in this chapter we shall discuss some of them more systematically. We shall find it convenient to change our notation, writing A_n instead of s_n, and A for the sum of the series instead of s. Thus $\sum a_n = A$ (C, 1) means

$$\lim \frac{A_0+A_1+...+A_n}{n+1} = A.$$

We shall also sometimes use A for the series, as well as for its sum, and say, for example, that 'A is summable (C, 1)' (naturally to sum A).

5.2. Hölder's means. The most obvious generalization is that first made by Hölder, who defined a sequence of methods which we shall call the (H, k) methods.

The (H, 1) method is the same as the (C, 1) method: thus

$$1-1+1-... = \tfrac{1}{2} \text{ (H, 1)}.$$

The method fails for $1-2+3-4+...$, since here the A_n are $1, -1, 2, -2, 3,...$, and

$$H_n^1 = \frac{A_0+A_1+...+A_n}{n+1}\dagger$$

is $\frac{1}{2}(n+2)/(n+1)$ if n is even and 0 if n is odd. We can, however, obtain a limit by repeating the averaging process; for the first of these values is $\frac{1}{2}+o(1)$, and so

$$H_n^2 = \frac{H_0^1+H_1^1+...+H_n^1}{n+1} \to \frac{1}{4}.$$

Similarly, three averagings will give $\frac{1}{8}$ as the sum of $1-3+6-10+...$.

We are thus led to define summability (H, k), for any positive integral k, as follows. We define H_n^k, for $k = 0, 1, 2,...$, by $H_n^0 = A_n$ and

$$H_n^{r+1} = \frac{H_0^r+H_1^r+...+H_n^r}{n+1}. \tag{5.2.1}$$

If $H_n^k \to A$ when $n \to \infty$, then we say that $\sum a_n$ is *summable* (H, k) *to sum* A, and write

$$a_0+a_1+a_2+... = A \text{ (H, } k). \tag{5.2.2}$$

By summability (H, 0) we mean convergence.

† We write H_n^1, H_n^2,... rather than $H_n^{(1)}$, $H_n^{(2)}$,... for convenience in printing: the indices cannot be read as powers.

5.3. Simple theorems concerning Hölder summability. We shall find that Hölder's definitions, although the most obvious generalizations of the (C, 1) definition, are for most purposes not the most convenient. They have, however, certain advantages. In particular, if we write $H_n^k(A)$ for the H_n^k formed from the partial sums A_n, and denote the sequence $(H_n^k(A))$ by $H^k(A)$, then it is obvious from the definitions that

$$H_n^k\{H^l(A)\} = H_n^l\{H^k(A)\} = H_n^{k+l}(A);$$

and this makes the proofs of some theorems particularly simple.

THEOREM 38. *If* $\sum a_n = A$ (H, k), *where* $k \geqslant 0$, *and* $k' > k$, *then* $\sum a_n = A$ (H, k').

This follows at once from the definitions and Cauchy's theorem of § 1.4.

THEOREM 39. *If* $\sum a_n = A$ (H, k), *then* $A_n = o(n^k)$ *and* $a_n = o(n^k)$.

For $H_n^k = A + o(1)$ and so

$$H_n^{k-1} = (n+1)H_n^k - nH_{n-1}^k = o(n), \quad H_n^{k-2} = (n+1)H_n^{k-1} - nH_{n-1}^{k-1} = o(n^2),$$

$$\ldots, \quad A_n = H_n^0 = (n+1)H_n^1 - nH_{n-1}^1 = o(n^k), \quad a_n = A_n - A_{n-1} = o(n^k).$$

This is the 'limitation theorem' for the (H, k) method. It shows, for example, that (as we saw directly in § 5.2) the series $1-2+3-4+\ldots$ cannot be summable (H, 1).

The next theorem reveals some of the inconveniences of the Hölder methods.

THEOREM 40. *The* (H, k) *method has the properties expressed by*

(α) $\sum Ca_n = C \sum a_n$, (β) $\sum (a_n + b_n) = \sum a_n + \sum b_n$,

(γ) $a_0 + a_1 + a_2 + \ldots = a_0 + (a_1 + a_2 + \ldots)$,

(δ) $a_0 + (a_1 + a_2 + \ldots) = a_0 + a_1 + a_2 + \ldots$.

Here each equation is to be interpreted in the sense 'if the right-hand side has a value, in the (H, k) sense, then the left-hand side has a value in the same sense, and the values are equal'. Thus (δ) means 'if $a_0 + a_1 + \ldots$ is summable to A, then $a_1 + a_2 + \ldots$ is summable to $A - a_0$'.

The properties (α) and (β) are trivial (and true of any linear method). If $k = 1$ and $b_n = a_{n+1}$, then $B_n = A_{n+1} - a_0$ and

$$\frac{B_0 + B_1 + \ldots + B_n}{n+1} = \frac{n+2}{n+1}\left(\frac{A_0 + A_1 + \ldots + A_{n+1}}{n+2} - a_0\right);$$

and (γ) and (δ) follow. But the relations between the means of the a_n and the b_n are not simple for higher values of k, and we postpone the rest of the proof to § 5.8.

5.4. Cesàro means. The Hölder means were defined by a process $\delta\sum.\delta\sum.\delta\sum...$, where $\sum$ is summation from 0 to n and δ is division by $n+1$, operating on $A_0, A_1,...$. The Cesàro means are averages defined by k summations followed by a single division.

We write

(5.4.1) $$A_n^0 = A_n = a_0+a_1+...+a_n,..., \quad A_n^k = A_0^{k-1}+A_1^{k-1}+...+A_n^{k-1},$$

and E_n^k for the value of A_n^k when $a_0 = 1$ and $a_n = 0$ for $n > 0$, i.e. when $A_n = 1$ for all n. If

(5.4.2) $$C_n^k(A) = A_n^k/E_n^k \to A$$

when $n \to \infty$, then we say that $\sum a_n$ is *summable* (C, k) *to sum* A, and write

(5.4.3) $$a_0+a_1+a_2+... = A \quad (\mathrm{C}, k).$$

It is easy to express A_n^k explicitly in terms of A_n or a_n. We have

$$(1-x)^{-p} = \sum (-1)^n \binom{-p}{n} x^n = \sum \binom{n+p-1}{p-1} x^n;$$

and

$$\sum A_n^k x^n = (1-x)^{-1} \sum A_n^{k-1} x^n = (1-x)^{-2} \sum A_n^{k-2} x^n ...$$
$$= (1-x)^{-k} \sum A_n x^n = (1-x)^{-k-1} \sum a_n x^n.$$

Thus $$\sum A_n^k x^n = \sum \binom{n+k-1}{k-1} x^n \sum A_n x^n$$

and

(5.4.4) $$A_n^k = \sum \binom{n-\nu+k-1}{k-1} A_\nu = \sum \binom{\nu+k-1}{k-1} A_{n-\nu}.\dagger$$

Similarly,

(5.4.5) $$A_n^k = \sum \binom{n-\nu+k}{k} a_\nu = \sum \binom{\nu+k}{k} a_{n-\nu}.$$

If $a_0 = 1$ and the remaining a_n are 0, then A_n^k reduces to $\binom{n+k}{k}$. Hence

(5.4.6) $$E_n^k = \binom{n+k}{k} = \frac{(k+1)(k+2)...(k+n)}{n!}.$$

Also $$\binom{n+k}{k} = \frac{(n+1)(n+2)...(n+k)}{k!} \sim \frac{n^k}{k!},$$

so that summability (C, k), to sum A, may also be defined by

(5.4.7) $$k!\,n^{-k} A_n^k \to A.$$

† Here we use a natural extension of the convention of § 3.1. A sum $\sum \alpha_\nu \beta_{n-\nu}$, without limits, is extended over those ν for which ν and $n-\nu$ are non-negative, i.e. over $0 \leqslant \nu \leqslant n$.

More generally, we have

$$\sum A_n^{k'} x^n = (1-x)^{-(k'-k)} \sum A_n^k x^n, \qquad \sum A_n^k x^n = (1-x)^{k'-k} \sum A_n^{k'} x^n;$$

and so

$$A_n^{k'} = \sum \binom{\nu+k'-k-1}{k'-k-1} A_{n-\nu}^k, \tag{5.4.8}$$

$$A_n^k = \sum (-1)^\nu \binom{k'-k}{\nu} A_{n-\nu}^{k'}. \tag{5.4.9}$$

These formulae are essentially the same, since

$$\binom{k'-k}{\nu} = (-1)^\nu \binom{\nu+k-k'-1}{k-k'-1},$$

so that (5.4.9) is (5.4.8) with k and k' interchanged; but the forms given are the most convenient when $k' > k$. Since the coefficient in (5.4.9) is 0 when (k' and k are integers and) $\nu > k'-k > 0$, it may also be written as

$$A_n^k = \sum_{\nu=0}^{k'-k} (-1)^\nu \binom{k'-k}{\nu} A_{n-\nu}^{k'}. \tag{5.4.10}$$

We can use (5.4.10) to define A_n^k for negative k. Thus, if $k = -p$ and $k' = 0$, it becomes

$$A_n^{-p} = A_n - \binom{p}{1} A_{n-1} + \binom{p}{2} A_{n-2} - \dots = (-1)^p \Delta^p A_{n-p}.$$

In particular

$$A_n^{-1} = -\Delta A_{n-1} = A_n - A_{n-1} = a_n; \tag{5.4.11}$$

and it is often convenient to use this convention.

5.5. Means of non-integral order. We have supposed so far (except in the last paragraph) that k is a positive integer, but the formulae (5.4.4)–(5.4.7) remain significant for non-integral k, and enable us to give more general definitions.

If k is a negative integer, and we define E_n^k either by (5.4.6), or as the coefficient of x^n in $(1-x)^{-k-1}$, then $E_n^k = 0$ for $n > -k-1$, and definition (5.4.2) fails. We must therefore exclude these values of k, and it proves best to suppose that $k > -1$. We then define A_n^k by (5.4.4) or (5.4.5), E_n^k by (5.4.6), and summability (C, k) by (5.4.2). The asymptotic formula for E_n^k is still valid if we interpret $k!$ as $\Gamma(k+1)$, and we can use (5.4.7) with this interpretation.

To show the desirability of the restriction $k > -1$, we suppose

$$\sum a_n x^n = (1-x)^{-p},$$

where p is positive and non-integral. Then

$$a_n = \frac{p(p+1)\dots(p+n-1)}{n!} \sim \frac{n^{p-1}}{\Gamma(p)},$$

so that $\sum a_n$ is a divergent series of positive terms. But $\sum A_n^k x^n = (1-x)^{-k-1-p}$, and in particular $\sum A_n^{-p-1} x^n = 1$. Hence $A_n^{-p-1} = 0$ for $n > 0$, and so (if we do not restrict k) $\sum a_n$ is summable $(C, -p-1)$ to sum 0. It would be very inconvenient for most purposes to attribute a finite sum to a divergent series of positive terms.†

We shall therefore suppose generally that $k > -1$; but it is sometimes convenient to use a special definition of summability $(C, -1)$. We shall say that $\sum a_n$ is summable $(C, -1)$ to sum A if (i) it converges to A and (ii) $a_n = o(n^{-1})$.

If $A_n^k = O(n^k)$ then we shall say that A_n *is bounded* (C, k), and write

$$A_n = O(1)\ (C, k).$$

More generally, by

$$A_n = o(n^l)\ (C,k), \qquad A_n = O(n^l)\ (C,k),$$

we shall mean

$$A_n^k = o(n^{l+k}), \qquad A_n^k = O(n^{l+k}).$$

And we shall use similar notations for other methods of summation: thus $\sum a_n = O(1)$ (A) will mean that $\sum a_n x^n = O(1)$ when $x \to 1-0$.

In what follows we shall sometimes work with a general k and sometimes restrict k to integral values. Most of the theorems with which we shall be concerned are true for all $k > -1$, but the proofs are often much simpler for integral k. Thus we have often to use the difference

$$\Delta^k u_n = u_n - \binom{k}{1} u_{n+1} + \binom{k}{2} u_{n+2} - \dots.$$

This is a finite sum when k is integral, but the generalization for non-integral k is an infinite series, and this often leads to serious complications. In such cases we shall usually suppose k integral.

5.6. A theorem concerning integral resultants.

The sum

$$c_n = \sum_{\mu+\nu=n} a_\mu b_\nu = \sum a_\nu b_{n-\nu} = \sum a_{n-\nu} b_\nu \tag{5.6.1}$$

and the integral

$$c(x) = \int a(t)b(x-t)\,dt = \int a(x-t)b(t)\,dt\ ‡ \tag{5.6.2}$$

are called the *resultants* of a_n, b_n and $a(x)$, $b(x)$. There is one pair of theorems concerning such resultants which we shall use repeatedly, and which will be particularly important in Ch. X.

THEOREM 41. *If* $r > -1$, $s > -1$, *and*

$$a_n \sim \binom{n+r}{r}\alpha \sim \frac{n^r}{\Gamma(r+1)}\alpha, \qquad b_n \sim \binom{n+s}{s}\beta \sim \frac{n^s}{\Gamma(s+1)}\beta, \tag{5.6.3}$$

then

$$c_n \sim \binom{n+r+s+1}{r+s+1}\alpha\beta \sim \frac{n^{r+s+1}}{\Gamma(r+s+2)}\alpha\beta. \tag{5.6.4}$$

† Though some definitions do this: thus $1+2+4+\dots = -1$ according to the $\mathfrak{E}$ definition of § 1.3. See also §§ 13.10 and 13.17.

‡ Here we use a convention similar to that of § 5.4: the range is $0 \leqslant t \leqslant x$. The German equivalent of resultant is *Faltung*.

THEOREM 42. *If $r > -1$, $s > -1$; $a(x)$ and $b(x)$ are integrable over any finite interval of positive x; and*

$$a(x) \sim \alpha x^r, \qquad b(x) \sim \beta x^s \tag{5.6.5}$$

when $x \to \infty$; then

$$c(x) \sim \frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)} \alpha\beta x^{r+s+1}. \tag{5.6.6}$$

In these theorems $a_n \sim \binom{n+r}{r}\alpha, \dots, a(x) \sim \alpha x^r, \dots$ are to be interpreted as $a_n = o(n^r), \dots, a(x) = o(x^r), \dots$ if α or β is 0; we leave the necessary modifications of the proofs to the reader. There are similar theorems in which hypotheses and conclusions involve O instead of o.

Theorem 41 may be deduced from Theorem 42 by taking $a(x) = a_n$ and $b(x) = b_n$ for $n \leqslant x < n+1$, when $c(n+1)$ reduces to c_n.

In proving Theorem 42 we may suppose $\alpha = \beta = 1$. We choose $\delta = \delta(\epsilon)$ so that

$$0 < \delta < \tfrac{1}{2}, \qquad \delta^{r+1} < (r+1)\epsilon, \qquad \delta^{s+1} < (s+1)\epsilon, \tag{5.6.7}$$

and also

$$\gamma = \frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)} = \int_0^1 u^r(1-u)^s\,du < \int_\delta^{1-\delta} u^r(1-u)^s\,du + \epsilon; \tag{5.6.8}$$

and write

$$c(x) = \int_0^{\delta x} + \int_{\delta x}^{(1-\delta)x} + \int_{(1-\delta)x}^{x} = c_1(x)+c_2(x)+c_3(x). \tag{5.6.9}$$

When δ is fixed we can choose $x_0 = x_0(\delta, \epsilon) = x_0(\epsilon)$ so that

$$(1-\epsilon)\int_{\delta x}^{(1-\delta)x} u^r(x-u)^s\,du < c_2(x) < (1+\epsilon)\int_{\delta x}^{(1-\delta)x} u^r(x-u)^s\,du,$$

$$(1-\epsilon)x^{r+s+1}\int_\delta^{1-\delta} u^r(1-u)^s\,du < c_2(x) < (1+\epsilon)x^{r+s+1}\int_\delta^{1-\delta} u^r(1-u)^s\,du,$$

for $x \geqslant x_0$. It follows, after (5.6.8), that

$$\overline{\lim}\,\frac{c_2(x)}{x^{r+s+1}} \leqslant (1+\epsilon)\int_\delta^{1-\delta} u^r(1-u)^s\,du \leqslant (1+\epsilon)\gamma,$$

$$\underline{\lim}\,\frac{c_2(x)}{x^{r+s+1}} \geqslant (1-\epsilon)\int_\delta^{1-\delta} u^r(1-u)^s\,du \geqslant (1-\epsilon)\gamma-\epsilon.$$

On the other hand, there are numbers H and K such that $|a(x)| < Kx^r$ and $|b(x)| < Kx^s$ for $x \geqslant H$. Hence if, as we may suppose, $\delta x_0 > H$ and $(1-\delta)x_0 > H$, we have

$$|c_1(x)| \leqslant Kx^s \int_0^{\delta x} |a(u)|\,du \leqslant Kx^s \int_0^H |a(u)|\,du + K^2x^s \int_H^{\delta x} u^r\,du$$

$$\leqslant Kx^s \int_0^H |a(u)|\,du + K^2 \frac{\delta^{r+1}}{r+1} x^{r+s+1}.$$

It follows from this and (5.6.7) that

$$\overline{\lim} \frac{|c_1(x)|}{x^{r+s+1}} \leqslant K^2 \frac{\delta^{r+1}}{r+1} < K^2\epsilon,$$

and there is plainly a similar inequality for $|c_3(x)|$.

Collecting our results we see that

$$\overline{\lim} \frac{c(x)}{x^{r+s+1}} \leqslant \overline{\lim} \frac{c_2(x)}{x^{r+s+1}} + \overline{\lim} \frac{|c_1(x)|}{x^{r+s+1}} + \overline{\lim} \frac{|c_3(x)|}{x^{r+s+1}} < (1+\epsilon)\gamma + 2K^2\epsilon,$$

$$\underline{\lim} \frac{c(x)}{x^{r+s+1}} \geqslant \underline{\lim} \frac{c_2(x)}{x^{r+s+1}} - \overline{\lim} \frac{|c_1(x)|}{x^{r+s+1}} - \overline{\lim} \frac{|c_3(x)|}{x^{r+s+1}} > (1-\epsilon)\gamma - (1+2K^2)\epsilon,$$

and so that $c(x) \sim \gamma x^{r+s+1}$.

We can, of course, prove Theorem 41 directly in a similar way, the part of the formula

$$\int u^r (x-u)^s\,du = \frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)} x^{r+s+1}$$

being played by an identity between binomial coefficients, viz.

$$(5.6.10) \qquad \sum \binom{\nu+r}{r}\binom{n-\nu+s}{s} = \binom{n+r+s+1}{r+s+1}.$$

5.7. Simple theorems concerning Cesàro summability. We begin by proving the theorems for Cesàro means corresponding to Theorems 38–40.

THEOREM 43. *If $k' > k > -1$ and $\sum a_n = A$ (C, k), then*

$$\sum a_n = A \quad (\mathrm{C}, k').$$

For, if $k' = k+\delta$, then

$$A_n^{k'} = \sum \binom{\nu+\delta-1}{\delta-1} A_{n-\nu}^k,$$

by (5.4.8); and $A_n^k \sim \binom{n+k}{k} A$. It follows from Theorem 41 that

$$A_n^{k'} \sim \binom{n+k+\delta}{k+\delta} A = \binom{n+k'}{k'} A.$$

In particular, taking $k = 0$, and writing k for k', we have

THEOREM 44. *If $k > 0$, then the* (C, k) *method is regular.*

Since the coefficients in $C_n^k(A)$ are non-negative when $k > 0$, it follows from Theorem 9 that the method is totally regular in the sense of § 3.6.

It is instructive to deduce Theorem 43 from Theorem 2. If we express $C_n^{k'}(A)$ in terms of $C_n^k(A)$ from (5.4.8), with $k' = k+\delta$ and $\nu, n-\nu$ interchanged, we find that

$$C_n^{k'}(A) = \sum c_{n,\nu} C_\nu^k(A),$$

where
$$c_{n,\nu} = \binom{n-\nu+\delta-1}{\delta-1}\binom{\nu+k}{k}\bigg/\binom{n+k+\delta}{k+\delta}$$

for $\nu \leqslant n$ and $c_{n,\nu} = 0$ for $\nu > n$. Then $c_{n,\nu} \geqslant 0$, $c_{n,\nu} = O(n^{-k-1})$ when ν is fixed and $n \to \infty$, and $\sum c_{n,\nu} = 1$ by (5.6.10), so that the conditions of Theorem 2 are satisfied. Also, since $c_{n,\nu} \geqslant 0$, $C_n^k(A) \to \infty$ implies $C_n^{k'}(A) \to \infty$.

Theorem 43 remains true for $k = -1$, if we use the definition of summability (C, -1) given in § 5.5, but it needs a different proof. Actually rather more is true.

THEOREM 45. *If $\sum a_n$ converges to A, and $a_n = O(n^{-1})$, then*

$$\sum a_n = A \quad (\mathrm{C}, -1+\delta)$$

for every positive δ.

We may suppose $A = 0$ and $\delta < 1$. We write

$$A_n^{-1+\delta} = \sum_{\nu=0}^{n} \binom{\nu+\delta-1}{\delta-1} a_{n-\nu} = \sum_{0}^{N-1} + \sum_{N}^{n} = S_1 + S_2, \tag{5.7.1}$$

where $N = [\varpi n]$, $0 < \varpi < 1$. Then

$$S_1 = O\left(\frac{1}{n}\right) + \sum_{1}^{N-1} O(\nu^{\delta-1}) O\left(\frac{1}{n-\nu}\right) = O\left(\frac{N^\delta}{n-N}\right) = O\left(\frac{\varpi^\delta}{1-\varpi} n^{\delta-1}\right),$$

uniformly in ϖ. We can therefore choose ϖ so that

$$|S_1| < \epsilon n^{\delta-1}. \tag{5.7.2}$$

Next, if $u_\nu = \binom{\nu+\delta-1}{\delta-1}$, $u_\nu - u_{\nu-1} = \binom{\nu+\delta-2}{\delta-2} = O(\nu^{\delta-2})$; and

$$\begin{aligned} S_2 &= u_N a_{n-N} + u_{N+1} a_{n-N-1} + \dots + u_n a_0 \\ &= A_0(u_n - u_{n-1}) + A_1(u_{n-1} - u_{n-2}) + \dots + A_{n-N-1}(u_{N+1} - u_N) + A_{n-N} u_N \\ &= \sum_{\nu=0}^{n-N-1} o(1) O(n^{\delta-2}) + o(1) O(n^{\delta-1}) = o(n^{\delta-1}). \end{aligned} \tag{5.7.3}$$

Finally, it follows from (5.7.1)–(5.7.3) that $n^{1-\delta} A_n^{-1+\delta} \to 0$, i.e. that $\sum a_n$ is summable (C, $-1+\delta$) to sum 0.

THEOREM 46. *If $k > -1$ and $\sum a_n = A$ (C, k), then $A_n^{k'} = o(n^k)$ for $k' < k$.*

This is the 'limitation theorem'. It is not necessary here to suppose $k' > -1$; in particular the result is true when $k' = -1$, $A_n^{k'} = a_n$.

We take $A = 0$, so that $A_n^k = o(n^k)$, and write $k' = k-\delta$, so that $\delta > 0$. Then, after (5.4.8) and (5.4.9),

$$A_n^{k'} = \sum \binom{\nu-\delta-1}{-\delta-1} A_{n-\nu}^k = \sum (-1)^\nu \binom{\delta}{\nu} A_{n-\nu}^k = \sum d_\nu A_{n-\nu}^k,$$

say. If δ is an integer then, by (5.4.10), $A_n^{k'}$ is a linear combination of $\delta+1$ of the A_n^k, with coefficients whose moduli are all less than $(1+1)^\delta = 2^\delta$; and so $A_n^{k'} = o(n^k)$.

If δ is not an integer, then $\Gamma(-\delta)d_n \sim n^{-\delta-1}$ and $\sum |d_n| < \infty$. We then write

$$A_n^{k'} = \left(\sum_0^{[\frac{1}{2}n]} + \sum_{[\frac{1}{2}n]+1}^{n}\right) d_\nu A_{n-\nu}^k = S_1+S_2.$$

Here

$$|S_1| \leqslant \sum_0^{[\frac{1}{2}n]} |d_\nu||A_{n-\nu}^k| = \sum_0^{[\frac{1}{2}n]} |d_\nu||o(n^k)| = o(n^k \sum |d_\nu|) = o(n^k),$$

$$|S_2| \leqslant \sum_{[\frac{1}{2}n]+1}^{n} |d_\nu||A_{n-\nu}^k| = O\Big(n^{-\delta-1} \sum_1^{n-[\frac{1}{2}n]} \mu^k\Big) = O(n^{k-\delta}) = o(n^k),$$

since $k > -1$ and $\delta > 0$; and the theorem follows.

THEOREM 47. *The* (C, k) *method has the properties* (α)—(δ) *of Theorem* 40.

It is only necessary to prove (γ) and (δ). We have to show that, if $b_n = a_{n+1}$, then either of $\sum a_n = A$ (C, k) and $\sum b_n = A-a_0$ (C, k) implies the other. But

$$\sum A_n^k x^n = (1-x)^{-k-1} \sum a_n x^n = (1-x)^{-k-1}(a_0+x\sum b_n x^n).$$

Hence $A_n^k = E_n^k a_0 + B_{n-1}^k$ for $n > 0$, and the conclusion follows.

THEOREM 48. *If* $\sum a_n$ *is summable* (C, k), *where* $k > -1$, *then*

$$a_m = (a_m - a_{m+1}) + (a_{m+1} - a_{m+2}) + \dots \quad (\mathrm{C}, k).$$

We may suppose, after Theorem 47, that $m = 0$. If $b_n = a_n - a_{n+1}$ then $B_n = a_0 - a_{n+1} = a_0 - u_n$, say, and

$$B_n^k = \sum \binom{n-\nu+k-1}{k-1}(a_0 - u_\nu) = \binom{n+k}{k} a_0 - U_n^{k-1}.$$

Now $\sum u_n$ is summable (C, k), by Theorem 47, and $U_n^{k-1} = o(n^k)$, by Theorem 46. Hence $B_n^k \sim \binom{n+k}{k} a_0$, and $\sum b_n$ is summable (C, k) to sum a_0.

The theorem may be stated in the form *if* $\sum a_n$ *is summable* (C, k) *then* $a_n \to 0$ (C, k), and is also true (and trivial) for $k = -1$.

5.8. The equivalence theorem. Our next theorem is distinctly more difficult.

THEOREM 49. *The* (C, k) *and* (H, k) *means are equivalent: if* $\sum a_n$ *is summable* (C, k), *then it is summable* (H, k) *to the same sum, and conversely.*

Here naturally k is integral, since we have defined Hölder means only for integral k. We begin by proving

THEOREM 50. *If*

$$m_n = \frac{s_0+s_1+\ldots+s_n}{n+1}, \tag{5.8.1}$$

then the hypotheses

$$s_n \to s \ (\mathrm{C}, k), \qquad m_n \to s \ (\mathrm{C}, k-1) \tag{5.8.2}$$

are equivalent.

We define s_n^k as we defined A_n^k in § 5.4, so that

$$s_n^k = \binom{n+k}{k} C_n^k(s); \tag{5.8.3}$$

and m_n^k, $C_n^k(m)$ similarly. We have, by partial summation,

$$\sum_{\nu=0}^{n} (\nu+p)u_\nu = (n+p)u_n^1 - \sum_{\nu=0}^{n-1} u_\nu^1 = (n+p+1)u_n^1 - u_n^2,$$

for any p and n. Hence, since $s_n^1 = (n+1)m_n$, we derive successively

$$s_n^2 = \sum_0^n (\nu+1)m_\nu = (n+2)m_n^1 - m_n^2, \qquad s_n^3 = (n+3)m_n^2 - 2m_n^3, \ldots,$$

$$s_n^k = (n+k)m_n^{k-1} - (k-1)m_n^k; \tag{5.8.4}$$

and from (5.8.3) and (5.8.4) it follows that

$$C_n^k(s) = kC_n^{k-1}(m) - (k-1)C_n^k(m). \tag{5.8.5}$$

First, $C_n^{k-1}(m) \to s$ implies $C_n^k(m) \to s$, and so $C_n^k(s) \to s$.

Secondly, suppose that $C_n^k(s) \to s$. Since $m_n^{k-1} = m_n^k - m_{n-1}^k$, (5.8.4) is

$$s_n^k = (n+1)m_n^k - (n+k)m_{n-1}^k$$

or

$$C_n^k(s) = (n+1)C_n^k(m) - nC_{n-1}^k(m). \tag{5.8.6}$$

From this it follows that

$$(n+1)C_n^k(m) = C_0^k(s) + C_1^k(s) + \ldots + C_n^k(s), \tag{5.8.7}$$

and therefore that $C_n^k(m) \to s$. Finally, (5.8.5) shows that $C_n^{k-1}(m) \to s$.

This proves Theorem 50: and it is easy to deduce Theorem 49. For,

applying Theorem 50 k times in succession, we see that the hypotheses
$$C_n^k(A) \to A, \quad C_n^{k-1}\{H^1(A)\} \to A, \quad \ldots, \quad C_n^1\{H^{k-1}(A)\} \to A, \quad H_n^k(A) \to A$$
are all equivalent.

It is plain that Theorem 40 (§ 5.3), the proof of which we postponed, now follows from Theorems 47 and 49.

5.9. Mercer's theorem and Schur's proof of the equivalence theorem. Schur's proof of Theorem 49 is similar in principle, but makes the relations between the various matrices involved more explicit. It depends on an important theorem of Mercer.

THEOREM 51. *If $\alpha > 0$ and*

$$t_n = \alpha s_n + (1-\alpha)m_n \to s, \tag{5.9.1}$$

then $s_n \to s$.

We define m_{-1} as 0. Then $s_n = (n+1)m_n - nm_{n-1}$ for $n = 0, 1, 2, \ldots$, and

$$t_n = (\alpha n+1)m_n - \alpha n m_{n-1} \quad (n = 0, 1, 2, \ldots). \tag{5.9.2}$$

We choose $q_0, q_1, q_2, \ldots$ so as to satisfy

$$q_0 = 1, \quad q_0 - \alpha q_1 = 0, \quad (\alpha+1)q_1 - 2\alpha q_2 = 0, \quad (2\alpha+1)q_2 - 3\alpha q_3 = 0, \quad \ldots.$$

Then

$$q_n = \frac{1}{\alpha}\cdot\frac{\alpha+1}{2\alpha}\cdot\frac{2\alpha+1}{3\alpha}\cdots\frac{(n-1)\alpha+1}{n\alpha} = \frac{\Gamma(n+\beta)}{\Gamma(\beta)\Gamma(n+1)} \sim \frac{n^{\beta-1}}{\Gamma(\beta)},$$

where $\beta = 1/\alpha$; and

$$q_0+q_1+\ldots+q_n \sim \frac{n^\beta}{\Gamma(\beta+1)} \sim (\alpha n+1)q_n. \tag{5.9.3}$$

Multiplying the equations (5.9.2) by $q_0, q_1, \ldots$, adding, and using (5.9.3) and Theorem 12, we obtain

$$m_n = \frac{q_0 t_0 + q_1 t_1 + \ldots + q_n t_n}{(\alpha n+1)q_n} \to s; \tag{5.9.4}$$

and it follows from (5.9.1) and (5.9.4) that $s_n \to s$. This proves Theorem 51.

We use the following notation. If the transformations T and U are the same, i.e. have the same matrices, we write $\mathrm{T} = \mathrm{U}$. If T and U have coefficients $c_{m,n}$ and $d_{m,n}$, then $\alpha\mathrm{T}+\beta\mathrm{U}$ is the transformation with coefficients $\alpha c_{m,n}+\beta d_{m,n}$. If $t = \mathrm{T}(s)$, as in § 3.1, and $u = \mathrm{U}(t)$, then we write

$$u = \mathrm{U}\{\mathrm{T}(s)\} = \mathrm{UT}(s).$$

If $\mathrm{UT} = \mathrm{TU}$, we say that T and U are *commutable*. We write T^2 for TT, T^3 for TT^2, and so on.

If T has an inverse, i.e. a transformation S such that $t = \mathrm{T}(s)$ implies $s = \mathrm{S}(t)$ and conversely, we write T^{-1} for S. If E is the identity, i.e. the transformation $t_m = s_m$, then $\mathrm{T}^{-1}\mathrm{T} = \mathrm{T}\mathrm{T}^{-1} = \mathrm{E}$. A triangular transformation in which $c_{m,m} \neq 0$ for all m has an inverse.

We write $\mathrm{H}^{(k)}$ and $\mathrm{C}^{(k)}$ for the (H, k) and (C, k) transformations, and H, C for $\mathrm{H}^{(1)}$, $\mathrm{C}^{(1)}$. Thus $\mathrm{H} = \mathrm{C}$ and $\mathrm{H}^{(k)} = \mathrm{H}^k$. If $(m+1)t_m = s_0 + \ldots + s_m$ then $s_m = (m+1)t_m - mt_{m-1}$. Thus the matrices of H and H^{-1} are

$$|\mathrm{H}| = \begin{pmatrix} 1 & 0 & 0 & 0 & . & . \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & . & . \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & . & . \\ . & . & . & . & . & . \end{pmatrix}, \qquad |\mathrm{H}^{-1}| = \begin{pmatrix} 1 & 0 & 0 & 0 & . & . \\ -1 & 2 & 0 & 0 & . & . \\ 0 & -2 & 3 & 0 & . & . \\ . & . & . & . & . & . \end{pmatrix}.$$

Since $A_n^{r-1} = A_n^r - A_{n-1}^r$ and $\binom{n+r}{r} C_n^r(A) = A_n^r$, we have

$$rC_n^{r-1}(A) = (n+r)C_n^r(A) - nC_{n-1}^r(A)$$
$$= \{(n+1)C_n^r(A) - nC_{n-1}^r(A)\} + (r-1)C_n^r(A);$$

and so
$$r\mathrm{C}^{(r-1)} = \mathrm{H}^{-1}\mathrm{C}^{(r)} + (r-1)\mathrm{C}^{(r)},$$

(5.9.5)
$$\mathrm{H}\mathrm{C}^{(r-1)} = \rho\mathrm{C}^{(r)} + (1-\rho)\mathrm{H}\mathrm{C}^{(r)} = \mathrm{S}^{(r)}\mathrm{C}^{(r)},$$

where $\rho = 1/r$ and $\mathrm{S}^{(r)}$ is the transformation

$$\mathrm{S}^{(r)} = \rho\mathrm{E} + (1-\rho)\mathrm{H}.$$

Hence $\mathrm{H}^{k-r+1}\mathrm{C}^{(r-1)} = \mathrm{H}^{k-r}\mathrm{S}^{(r)}\mathrm{C}^{(r)}$ for $0 < r \leqslant k$. But H^{k-r} is commutable with H and with E, and so with $\mathrm{S}^{(r)}$; and therefore

(5.9.6)
$$\mathrm{H}^{k-r+1}\mathrm{C}^{(r-1)} = \mathrm{S}^{(r)}\mathrm{H}^{k-r}\mathrm{C}^{(r)} \quad (0 < r \leqslant k).$$

We define $\mathrm{T}^{(r)}$ by

(5.9.7)
$$\mathrm{T}^{(r)} = \mathrm{H}^{k-r}\mathrm{C}^{(r)} \quad (0 \leqslant r \leqslant k),$$

so that $\mathrm{T}^{(k)} = \mathrm{C}^{(k)}$ and $\mathrm{T}^{(0)} = \mathrm{H}^k$. Then (5.9.6) is $\mathrm{T}^{(r-1)} = \mathrm{S}^{(r)}\mathrm{T}^{(r)}$; and therefore

(5.9.8)
$$t_n^{(r-1)} = \frac{1}{r}t_n^{(r)} + \left(1 - \frac{1}{r}\right)\frac{t_0^{(r)} + t_1^{(r)} + \ldots + t_n^{(r)}}{n+1},$$

$t_n^{(r)}$ being the result of operating on A_n with $\mathrm{T}^{(r)}$. Hence, by Theorem 51, the hypotheses $t_n^{(r)} \to A$ and $t_n^{(r-1)} \to A$ are equivalent. That is to say, $\mathrm{T}^{(r)}$ and $\mathrm{T}^{(r-1)}$ are equivalent, and therefore $\mathrm{T}^{(k)}$ and $\mathrm{T}^{(0)}$ are equivalent.

It will be observed that here we use the transformations $\mathrm{C}^{(k)}, \mathrm{H}\mathrm{C}^{(k-1)}, \mathrm{H}^2\mathrm{C}^{(k-2)}, \ldots, \mathrm{H}^k$, whereas in § 5.8 we used $\mathrm{C}^{(k)}, \mathrm{C}^{(k-1)}\mathrm{H}, \mathrm{C}^{(k-2)}\mathrm{H}^2, \ldots, \mathrm{H}^k$. Actually H^p is commutable with $\mathrm{C}^{(q)}$ for all p and q, so that $\mathrm{H}^r\mathrm{C}^{(k-r)} = \mathrm{C}^{(k-r)}\mathrm{H}^r$, and the two sets of transformations are the same. This is not difficult to prove directly, but the full reason for it will not appear until §§ 11.3–4.

5.10. Other proofs of Mercer's theorem. From the many other proofs of Theorem 51 we select two.

(A) *Knopp's proof.* One proof, due to Knopp, has the merit of avoiding all algebraical calculations. We may suppose without loss of generality that s_n is real; and it is sufficient to show that s_n tends to a limit.

Given $n > 0$, we distinguish the two cases (*a*) $s_n < m_n$, (*b*) $s_n \geqslant m_n$. Since $s_n = (n+1)m_n - nm_{n-1}$, $s_n < m_n$ implies $m_{n-1} > m_n$ (and $s_n > m_n$, $s_n = m_n$ imply $m_{n-1} < m_n$ and $m_{n-1} = m_n$ respectively).

(i) Suppose that

$$\overline{\lim}\, m_n = \infty. \tag{5.10.1}$$

Then, given G, there is a p for which $m_p > G$. If $n = p$ is in case (*a*), then $m_{p-1} > m_p > G$. If also $p-1$ is in case (*a*), then $m_{p-2} > m_{p-1} > G$; and so on. If all of $p, p-1, \ldots, 2$ are in case (*a*), then $m_1 > G$, and this is impossible for large G. Hence one of these numbers must be in case (*b*), and there is a q such that $s_q \geqslant m_q > G$. But then

$$t_q = \alpha s_q + (1-\alpha)m_q = m_q + \alpha(s_q - m_q) > G,$$

a contradiction for large G, since t_n is bounded. Hence $\overline{\lim}\, m_n$ is finite; and similarly $\underline{\lim}\, m_n$ is finite, so that m_n is bounded.

(ii) Suppose that (m_n is bounded and)

$$l = \underline{\lim}\, m_n < \overline{\lim}\, m_n = L. \tag{5.10.2}$$

Then there are numbers h and H such that $h < H$ and each of $m_n < h$, $m_n > H$ is true for an infinity of n. Suppose, for example, that

$$m_p < h, \qquad m_q > H, \qquad q > p. \tag{5.10.3}$$

If q is in case (*a*) then, as before, $m_{q-1} > m_q > H$. If all of $q, q-1, \ldots, p+1$ are in case (*a*), then $m_p > H > h$, in contradiction to (5.10.3). It follows that there is an r, greater than p, for which $s_r \geqslant m_r > H$ and

$$t_r = \alpha s_r + (1-\alpha)m_r = m_r + \alpha(s_r - m_r) > H. \tag{5.10.4}$$

And, since p may be as large as we please, (5.10.4) is true for an infinity of r. Similarly $t_r < h$ for an infinity of r; and this and (5.10.4) together contradict the hypothesis that t_r tends to a limit. It follows that (5.10.2) is false, and that m_n tends to a limit; and therefore, by (5.9.1), s_n tends to a limit.

(B) *Hardy's proof.* Another proof, by Hardy, gives rather more, in particular the extension of the theorem to complex α. It is convenient to begin by a trivial transformation of the theorem.

We write $\qquad u_{n+1} = \alpha(s_0 + s_1 + \ldots + s_n), \qquad a = (\alpha - 1)/\alpha.$

Then (5.9.1), with $n-1$ for n, becomes

$$u_n - u_{n-1} - au_n/n \to s, \tag{5.10.5}$$

and positive values of α correspond to values of a less than 1. Mercer's theorem asserts that $u_n - u_{n-1}$ and u_n/n then tend to $s/(1-a)$. We prove, more generally,

THEOREM 52. *If $a = \alpha + i\beta$ and $\alpha \neq 1$, then* (5.10.5) *implies*

$$u_n = C\frac{\Gamma(n+1)}{\Gamma(n+1-a)} + \frac{sn}{1-a} + o(n).$$

If $\alpha < 1$, then $C = 0$.

We may suppose that $s = 0$. We write

$$u_n = \frac{\Gamma(n+1)}{\Gamma(n+1-a)}\phi_n = f_n\phi_n.$$

Then $f_n \sim n^a$ and $f_n - f_{n-1} = af_n/n$. Hence

$$f_{n-1}(\phi_n-\phi_{n-1}) = f_n\phi_n - f_{n-1}\phi_{n-1} - (f_n - f_{n-1})\phi_n$$
$$= u_n - u_{n-1} - au_n/n = o(1),$$

and so $\phi_n - \phi_{n-1} = o(n^{-\alpha})$. If $\alpha < 1$ then

$$\phi_n = \phi_0 + \sum_1^n (\phi_m - \phi_{m-1}) = \phi_0 + \sum_1^n o(m^{-\alpha}) = o(n^{1-\alpha})$$

and $u_n = O(n^\alpha)o(n^{1-\alpha}) = o(n)$. If $\alpha > 1$ then the series $\sum(\phi_m - \phi_{m-1})$ is convergent, ϕ_n tends to a limit C,

$$\phi_n = C - \sum_n^\infty (\phi_{m+1} - \phi_m) = C + \sum_n^\infty o(m^{-\alpha}) = C + o(n^{1-\alpha}),$$

and $u_n = Cf_n + o(n)$.

There are theorems of the same kind concerning 'asymptotic differential equations', and one particularly simple theorem which we shall use later, viz.

THEOREM 53. *If $f(x)+f'(x) \to 0$ when $x \to \infty$, then $f(x) \to 0$.*

This may be proved directly as follows. If f' is of fixed sign from a certain x onwards, then f is monotonic. Thus f tends to a (possibly infinite) limit l, and $f' \to -l$; and these conclusions are contradictory unless $l = 0$. If, on the other hand, f' assumes values of either sign for values of x beyond all limit, then $f \to 0$ when $x \to \infty$ through the values which make f a maximum or minimum, and therefore when $x \to \infty$ in any manner.

5.11. Infinite limits. It is natural to ask whether the equivalence theorem extends to the case in which the limits are infinite. Here the answer is negative.

THEOREM 54. *If $s_n \to \infty$ (C, k) then $s_n \to \infty$ (H, k). The converse is false when $k > 1$.*

Here again k is integral. It follows from (5.9.5) that

$$\mathrm{H} = \mathrm{C}^{(1)}, \qquad \mathrm{H}^2 = \mathrm{HC}^{(1)} = \tfrac{1}{2}\mathrm{C}^{(2)} + \tfrac{1}{2}\mathrm{HC}^{(2)},$$
$$\mathrm{H}^3 = \tfrac{1}{2}\mathrm{HC}^{(2)} + \tfrac{1}{2}\mathrm{H}^2\mathrm{C}^{(2)} = \tfrac{1}{3}\mathrm{C}^{(3)} + \tfrac{1}{2}\mathrm{HC}^{(3)} + \tfrac{1}{6}\mathrm{H}^2\mathrm{C}^{(3)},$$

and generally

(5.11.1) $$\mathrm{H}^k = \sum_{p=0}^{k-1} a_{k,p}\mathrm{H}^p\mathrm{C}^{(k)},$$

where $a_{k,p} > 0$. Hence

(5.11.2) $$H_n^k(s) = \sum_{p=0}^{k-1} a_{k,p} H_n^p\{C^{(k)}(s)\}.$$

Also

(5.11.3) $$H_n^p\{C^{(k)}(s)\} = \sum_{q=0}^n h_{n,p,q}C_q^{(k)}(s),$$

where $h_{n,p,q} > 0$; and it follows from (5.11.2) and (5.11.3) that

(5.11.4) $$H_n^k(s) = \sum_{q=0}^n b_{k,n,q}C_q^{(k)}(s),$$

where
$$b_{k,n,q} = \sum_{p=0}^{k-1} a_{k,p}h_{n,p,q} > 0.$$

Consideration of the case in which $s_n = 1$ for all n shows that $\sum b_{k,n,q} = 1$.

The equivalence theorem shows that the transformation (5.11.4) is regular, and, since $b_{k,n,q} > 0$, it satisfies the conditions of Theorem 9. Hence

$$\underline{\lim}\, C_n^{(k)}(s) \leqslant \underline{\lim}\, H_n^k(s) \leqslant \overline{\lim}\, H_n^k(s) \leqslant \overline{\lim}\, C_n^{(k)}(s), \tag{5.11.5}$$

and $s_n \to \infty$ (C, k) implies $s_n \to \infty$ (H, k). This proves the positive half of Theorem 54.

To prove the negative half, suppose that $k > 1$ and $C_{2m}^{(k)}(s) = 2m$, $C_{2m+1}^{(k)}(s) = 0$. These equations define a sequence (s_n) for which

$$\underline{\lim}\, C_n^{(k)}(s) = 0, \qquad \overline{\lim}\, C_n^{(k)}(s) = \infty,$$

and $H_n\{C^{(k)}(s)\} \to \infty$. By (5.11.2), $H_n^k(s) > a_{k,1} H_n\{C^{(k)}(s)\} \to \infty$. Thus $s_n \to \infty$ (H, k), but $s_n \to \infty$ (C, k) is false.

5.12. Cesàro and Abel summability. Theorem 43 shows that the strength of the (C, k) methods increases with k. Our next theorems show that the A method is stronger than any of them.

THEOREM 55. *If* $\sum a_n = A$ (C, k), *for some* k, *then* $\sum a_n = A$ (A).

THEOREM 56. *There are series summable* (A) *but not summable* (C, k) *for any* k.

We need a lemma, important in itself.

THEOREM 57. *If* $d_n > 0$, $\sum d_n = \infty$, $\sum d_n x^n$ *is convergent for* $0 \leqslant x < 1$, *and* $c_n \sim A d_n$, *where* $A \neq 0$, *then*

$$C(x) = \sum c_n x^n \sim AD(x) = A \sum d_n x^n$$

when $x \to 1$.

We may suppose c_n real and $A = 1$. Then c_n/d_n lies between $1-\epsilon$ and $1+\epsilon$ for $n > N = N(\epsilon)$. Hence, on the one hand,

$$C(x) = \sum_0^N c_n x^n + \sum_{N+1}^\infty c_n x^n \leqslant (1+\epsilon)D(x) + \sum_0^N |c_n| x^n,$$

and on the other

$$C(x) \geqslant (1-\epsilon)D(x) - \sum_0^N d_n x^n - \sum_0^N |c_n| x^n.$$

Since

$$\lim D(x) \geqslant \lim \sum_0^N d_n x^n = D_N$$

for every N, and so $D(x) \to \infty$, it follows that

$$\overline{\lim_{x\to 1}} \frac{C(x)}{D(x)} \leqslant 1+\epsilon, \qquad \underline{\lim_{x\to 1}} \frac{C(x)}{D(x)} \geqslant 1-\epsilon,$$

and therefore that $C(x) \sim D(x)$.

Theorem 55 is a corollary. We may suppose $A \neq 0$. Then, as in § 5.4,

$$f(x) = \sum a_n x^n = \frac{\sum A_n^k x^n}{(1-x)^{-k-1}} = \frac{\sum A_n^k x^n}{\sum E_n^k x^n};$$

and $A_n^k \sim AE_n^k$, so that $f(x) \to A$.

To prove Theorem 56, we define a_n by

$$f(x) = e^{1/(1+x)} = \sum a_n x^n. \tag{5.12.1}$$

Then $f(x)$ is regular except for $x = -1$, so that the series is convergent for $|x| < 1$; and $f(x) \to e^{\frac{1}{2}}$ when $x \to 1$. On the other hand, a_n is not $O(n^k)$ for any k; for this would involve

$$f(x) = O(\sum n^k |x|^n) = O\{(1-|x|)^{-k-1}\},$$

uniformly in the circle $|x| < 1$, whereas $f(x)$ tends to infinity like $e^{1/(1-|x|)}$ when $x \to -1$ by real values. It follows from Theorem 46 that $\sum a_n$ is not summable (C, k) for any k.

A more elegant example of a series with the properties desired is $\sum (-1)^n e^{c\sqrt{n}}$, where $c > 0$. The a_n of (5.12.1) is roughly of this type, but the proof of this is more troublesome.

5.13. Cesàro means as Nörlund means. The (C, k) means are the (N, p_n) means with

$$p_n = \binom{n+k-1}{k-1}, \qquad p(x) = \sum \binom{n+k-1}{k-1} x^n = (1-x)^{-k}.$$

The (H, k) means are not Nörlund means (except when $k = 1$). It is interesting to find examples of Nörlund means (*a*) stronger than any Cesàro mean and (*b*) weaker than any Cesàro mean of positive order.

(*a*) We suppose k integral, and take

$$p_n = \binom{n+k-1}{k-1}, \qquad q_n = e^{\sqrt{n}},$$

when $P_n = O(n^k)$, $Q_n \sim 2\sqrt{n}\, e^{\sqrt{n}}$; and define κ_n by

$$\kappa(x) = \sum \kappa_n x^n = \frac{q(x)}{p(x)} = (1-x)^k q(x),\dagger$$

so that

$$\kappa_n = (-1)^k \Delta^k e^{\sqrt{(n-k)}} \sim 2^{-k} n^{-\frac{1}{2}k} e^{\sqrt{n}}$$

for $n \geqslant k$. We have to show that summability (N, p_n) implies summability (N, q_n), and we use Theorem 19. The second condition of the theorem is plainly satisfied, and it is enough to prove that

$$\sum e^{\sqrt{(n-m)}} (n-m)^{-\frac{1}{2}k} m^k = O(\sqrt{n}\, e^{\sqrt{n}}),$$

the summation extending over $0 < m < n$. The terms in which $m > \frac{1}{2}n$ give $O(e^{c\sqrt{n}})$ with a $c < 1$. Finally,

$$\sqrt{n} - \sqrt{(n-m)} > \frac{m}{2\sqrt{n}}, \qquad e^{\sqrt{(n-m)}} < e^{\sqrt{n}} e^{-\frac{1}{2}m/\sqrt{n}},$$

† We use κ for the k of § 4.3, since k is required otherwise.

and the remaining terms give

$$O(e^{\sqrt{n}}n^{-\frac{1}{2}k}\sum m^k e^{-\frac{1}{2}m/\sqrt{n}}) = O\{e^{\sqrt{n}}n^{-\frac{1}{2}k}(1-e^{-\frac{1}{2}n^{-\frac{1}{2}}})^{-k-1}\}$$
$$= O\{e^{\sqrt{n}}n^{-\frac{1}{2}k}n^{\frac{1}{2}(k+1)}\} = O(\sqrt{n}\,e^{\sqrt{n}}).$$

(*b*) The means for which $p_n = (n+1)^{-1}$ have been called by M. Riesz 'harmonic' means. We take $q_n = \binom{n+k-1}{k-1}$, where $k < 1$, so that (N, q_n) is (C, k). Then $p_n^2 < p_{n-1}p_{n+1}$, and $p_n/p_{n-1} < q_n/q_{n-1}$ if $(n+1)(n+k-1) > n^2$, i.e. if $n > (1-k)/k$. Thus the conditions of Theorem 23 are satisfied, and summability (N, p_n) implies summability (C, k) for every positive k.

5.14. Integrals. The definitions for integrals corresponding to those of §§ 5.2–5 are as follows. We take the lower limit of integration to be 0, and suppose, to avoid minor complications, that $a(x)$ is bounded in every finite interval $(0, X)$.†

We write

$$H^0(x) = A(x) = \int_0^x a(t)\,dt, \qquad H^k(x) = \frac{1}{x}\int_0^x H^{k-1}(t)\,dt.$$

If $H^k(x) \to A$ when $x \to \infty$, we write

$$A(x) \to A \quad (\mathrm{H}, k), \qquad \int a(x)\,dx = A \quad (\mathrm{H}, k),‡$$

and say that the integral is summable (H, k) to A. If

$$A_0(x) = A(x), \qquad A_k(x) = \int_0^x A_{k-1}(t)\,dt,$$

and
$$k!\,x^{-k}A_k(x) \to A,$$
then we write

$$A(x) \to A \quad (\mathrm{C}, k), \qquad \int a(x)\,dx = A \quad (\mathrm{C}, k),$$

and say that the integral is summable (C, k) to A.

These definitions are for integral k. If k is integral, then

$$A_k(x) = \int_0^x A_{k-1}(t)\,dt = \int_0^x (x-t)A_{k-2}(t)\,dt \ldots \tag{5.14.1}$$
$$= \frac{1}{(k-1)!}\int_0^x (x-t)^{k-1}A(t)\,dt = \frac{1}{k!}\int_0^x (x-t)^k a(t)\,dt,$$

by repeated partial integration; and these formulae suggest the exten-

† See the note at the end of the chapter.

‡ Integrals without limits being as usual over $(0, \infty)$.

sion of the definitions to non-integral k. We say that the integral $\int a(x)\,dx$ is summable (C,k), where $k>0$, to sum A, if

(5.14.2)
$$\frac{\Gamma(k+1)}{x^k}A_k(x)=\frac{k}{x^k}\int_0^x (x-t)^{k-1}A(t)\,dt=\int_0^x\left(1-\frac{t}{x}\right)^k a(t)\,dt\to A.$$

The second form, with $a(t)$, may be used for all $k>-1$.

If $A_k(x)$ is defined by (5.14.2), and $k>-1$, $l>0$, then

$$\frac{1}{\Gamma(l)}\int_0^x (x-t)^{l-1}A_k(t)\,dt=\frac{1}{\Gamma(k+1)\Gamma(l)}\int_0^x (x-t)^{l-1}\,dt\int_0^t (t-u)^k a(u)\,du$$

$$=\frac{1}{\Gamma(k+1)\Gamma(l)}\int_0^x a(u)\,du\int_u^x (t-u)^k(x-t)^{l-1}\,dt$$

$$=\frac{1}{\Gamma(k+l+1)}\int_0^x (x-u)^{k+l}a(u)\,du.$$

Thus

(5.14.3)
$$A_{k+l}(x)=\frac{1}{\Gamma(l)}\int_0^x (x-t)^{l-1}A_k(t)\,dt\quad(k>-1,l>0).$$

This is the analogue of (5.4.8).

5.15. Theorems concerning summable integrals. There are theorems for integrals corresponding to most of those of §§ 5.3–11, and the proofs are usually a little simpler than those of the theorems for series. There is, however, one important difference. If $\sum a_n$ is convergent then $a_n\to 0$, whereas there is no corresponding theorem for integrals. Thus there is no limitation theorem such as Theorem 46, and this destroys the analogy in some ways.

We summarize the main results, leaving the proofs, for the most part, to the reader, and emphasizing only what points of difference there are.

If $\int a(x)\,dx$ is summable (C,k), where $k>-1$, then it is summable (C,k') for $k'>k$. The proof depends on (5.14.3), and is otherwise similar to that of Theorem 43.

The methods have the properties analogous to those of Theorem 40. In particular

(5.15.1)
$$\int_0^\infty a(x)\,dx=\int_0^c a(x)\,dx+\int_c^\infty a(x)\,dx\quad(C,k)$$

if the last integral is defined as $\int a(c+y)\,dy$ and either side of the equation has a meaning.

The (H, k) and (C, k) definitions are equivalent. This is most easily proved by a modification of the proof of § 5.8. We have to show that, if $s_k(x)$ is defined like $A_k(x)$ in § 5.14, and

$$C^k(x,s) = k!\,x^{-k}s_k(x), \qquad m(x) = \frac{1}{x}\int_0^x s(t)\,dt,$$

then the assertions

$$(5.15.2) \qquad C^k(x,s) \to s, \qquad C^{k-1}(x,m) \to s$$

are equivalent. Now $s_1(x) = xm(x)$, $s_2(x) = xm_1(x)-m_2(x)$,..., where

$$m_1(x) = \int_0^x m(t)\,dt, \quad m_2(x) = \int_0^x m_1(t)\,dt,\ldots,$$

and generally

$$(5.15.3) \qquad s_k(x) = xm_{k-1}(x)-(k-1)m_k(x),$$

by repeated partial integration; and this is equivalent to

$$(5.15.4) \qquad C^k(x,s) = kC^{k-1}(x,m)-(k-1)C^k(x,m).$$

From this it follows that the second of (5.15.2) implies the first.

Next, (5.15.3) gives

$$\frac{s_k(x)}{x^k} = \frac{d}{dx}\left(\frac{m_k(x)}{x^{k-1}}\right), \qquad \frac{m_k(x)}{x^{k-1}} = \int_0^x \frac{s_k(t)}{t^k}\,dt,$$

and so

$$C^k(x,m) = \frac{1}{x}\int_0^x C^k(t,s)\,dt.$$

Hence $C^k(x,s) \to s$ implies $C^k(x,m) \to s$, and so, after (5.15.4),

$$C^{k-1}(x,m) \to s.$$

This proves the equivalence of the two assertions (5.15.2), and the proof of the main theorem then follows as in § 5.8.

5.16. Riesz's arithmetic means. The formulae (5.14.2) suggest a modification of the definitions of §§ 5.4–5. If k is integral then, in the notation of § 5.4,

$$C_n^k(A) = \binom{n+k}{k}^{-1}\sum_{\nu=0}^{n}\binom{n-\nu+k}{k}a_\nu$$

$$= \sum_{\nu=0}^{n}\left(1-\frac{\nu}{n+1}\right)\left(1-\frac{\nu}{n+2}\right)\cdots\left(1-\frac{\nu}{n+k}\right)a_\nu.$$

If here we replace all of $n+1, n+2,..., n+k$ by n, we obtain a new mean

(5.16.1) $$R_n^k(A) = \sum \left(1-\frac{\nu}{n}\right)^k a_\nu$$

more strictly analogous to the integral mean (5.14.2); and this led M. Riesz to suggest

(5.16.2) $$R_n^k(A) \to A$$

as a new definition. We may plainly allow k to have any positive value, but negative k are inadmissible.

Riesz found, however, that this definition did not lead to satisfactory results; the means $R_n^k(A)$ have, for the larger values of k, properties quite unlike those of the corresponding Cesàro means. He was therefore led to modify the definition by the introduction of a continuous parameter ω. The means thus obtained are the typical means of § 4.16, with $\lambda_n = n$.

We write

(5.16.3) $$R^k(\omega) = R^k(\omega, A) = \frac{T^k(\omega)}{\omega^k} = \sum_{\nu \leqslant \omega} \left(1-\frac{\nu}{\omega}\right)^k a_\nu,$$

where $k > 0$. If $R^k(\omega) \to A$ when $\omega \to \infty$, then we say that $\sum a_n$ is summable (R, n, k) to sum A. We then find that summability (R, n, k) is equivalent to summability (C, k). We confine our attention to integral k, the proof for general k being rather troublesome.

THEOREM 58. *If k is integral, and $\sum a_n$ is summable* (C, k), *then it is summable* (R, n, k) *to the same sum; and conversely.*

We may suppose the sum zero. We have then to show that the hypotheses

(5.16.4) $A_n^k = o(n^k)$, (5.16.5) $T^k(\omega) = o(\omega^k)$

are equivalent. We suppose that $\omega = n+\theta$, where n is an integer and $0 \leqslant \theta < 1$.

(i) Assume (5.16.4). Since $T^k(\omega) = \sum (n-\nu+\theta)^k a_\nu$, we have

$$\sum T^k(\omega)x^n = \sum (n+\theta)^k x^n \sum a_n x^n = g(x,\theta) \sum A_n^k x^n,$$

where

$$g(x,\theta) = (1-x)^{k+1} \sum (n+\theta)^k x^n = (1-x)^{k+1} x^{-\theta} \left(x\frac{d}{dx}\right)^k \frac{x^\theta}{1-x} = \sum_{j=0}^{k} c_j(\theta) x^j,$$

and the coefficients $c_j(\theta)$ are polynomials in θ of degree k. Hence

$$\sum_{n=0}^{\infty} T^k(\omega) x^n = \sum_{j=0}^{k} c_j(\theta) x^j \sum_{n=0}^{\infty} A_n^k x^n, \qquad T^k(\omega) = \sum_{\nu=0}^{k} c_\nu(\theta) A_{n-\nu}^k,$$

and (5.16.5) follows, with the necessary uniformity in θ.

(ii) Assume (5.16.5), and suppose that $0 < \theta_0 < \theta_1 \dots < \theta_k < 1$. Then we can determine $q_0, q_1, \dots, q_k$ so that

$$\binom{n+k}{k} = \sum_{r=0}^{k} q_r (n+\theta_r)^k$$

identically. For, if we equate the coefficients of different powers of n, we obtain a system of equations $\sum \theta_r^j q_r = C_j$, where $j = 0, 1, \dots, k$, in the q_r, and the determinant of the coefficients is

$$|\theta_r^j| = \prod_{l>m} (\theta_l - \theta_m) \neq 0.$$

We then have

$$A_n^k = \sum_{\nu=0}^{n} \binom{n-\nu+k}{k} a_\nu = \sum_{r=0}^{k} q_r T^k(n+\theta_r),$$

and (5.16.4) follows.

We add a few remarks to show the inadequacy of the definition (5.16.2). When $k = 1$,

$$C_n^1(A) = \frac{1}{n+1} \sum (n+1-\nu) a_\nu = R_{n+1}^1(A),$$

so that the definition is equivalent to Cesàro's; but there is no such equivalence for larger k. Suppose, for example, that $k = 2$. Then

$$\sum (n+1)^2 R_{n+1}^2(A) x^n = \sum (n+1)^2 x^n \sum a_n x^n = \frac{1+x}{(1-x)^3} \sum a_n x^n.$$

If we define a_n by

$$\sum a_n x^n = (1-x)(1+x)^{-3} = \sum (-1)^n (n+1)^2 x^n,$$

then a_n is of order n^2, and so $\sum a_n$ is not summable (C, 2); but

$$\sum (n+1)^2 R_{n+1}^2(A) x^n = (1-x^2)^{-2} = 1 + 2x^2 + 3x^4 + \dots$$

and $R_n^2(A) = O(n^{-1}) = o(1)$.

When $k = 3$, (5.16.2) does not imply the summability (C, k) of the series for *any* k, or even its summability (A). For

$$\sum (n+1)^3 x^n = (1-x)^{-4}(1+4x+x^2)$$

has a zero at $x = -2+\sqrt{3} = \alpha$, inside the unit circle. If we define a_n by $\sum a_n x^n = (1-x)/(\alpha - x)$, then $R_n^3(A) = o(1)$, but $\sum a_n x^n$ is convergent only for $|x| < \alpha < 1$.

It is instructive to consider this question in the light of §§ 4.3–4. The (C, 2) means are the (N, q_n) means with $q(x) = \sum q_n x^n = (1-x)^{-2}$; and

$$T^2(n+1) = (n+1)^2 R_{n+1}^2(A)$$

is the coefficient of x^n in

$$\sum (n+1)^2 x^n \sum a_n x^n = \sum p_n x^n \sum A_n x^n,$$

where $\quad p_n = (n+1)^2 - n^2 = 2n+1, \qquad p(x) = \sum p_n x^n = \dfrac{1+x}{(1-x)^2},$

so that $R_{n+1}^2(A)$ is the (N, p_n) mean for this p_n. In the notation of § 4.3 we have $k(x) = (1+x)^{-1} = 1 - x + x^2 - \dots$, so that $\sum |k_n| = \infty$. The equivalence is destroyed by the zero of $p(x)$ at $x = -1$, and it naturally fails more completely when $p(x)$ has a zero with $|x| < 1$.

5.17. Uniformly distributed sequences. We end this chapter by a short excursion into a different field.

We suppose that $0 \leqslant s_n \leqslant 1$ for every n, and denote the interval $0 \leqslant a \leqslant x \leqslant b \leqslant 1$ by I. If n_I is the number of $s_0, s_1, \ldots, s_n$ which fall in I, and $n_I \sim nI$ when $n \to \infty$,† for every I, then we say that the sequence (s_n) is *uniformly distributed* in $(0, 1)$.

We denote the characteristic function of I, 1 in I and 0 elsewhere, by $I(x)$. If $f(x) = I(x)$ then

$$\frac{f(s_0)+f(s_1)+\ldots+f(s_n)}{n+1} = \frac{n_I}{n+1}, \qquad \int f(x)\,dx = I.\dagger$$

Thus the assertion of uniform distribution is equivalent to the assertion that

$$f(s_n) \to \int f(x)\,dx \quad (\mathrm{C}, 1) \tag{5.17.1}$$

for every $f(x) = I(x)$. We now prove

THEOREM 59. *If (s_n) is uniformly distributed, then* (5.17.1) *is true for every Riemann integrable $f(x)$.*

We may plainly suppose f real. If (s_n) is uniformly distributed, then (5.17.1) is true for $f(x) = I(x)$. It follows by multiplication and addition that it is true for any finite step-function. If f is Riemann integrable, then there are finite step-functions f_1 and f_2 such that $f_1 \leqslant f \leqslant f_2$ and

$$0 \leqslant \int f_2\,dx - \int f_1\,dx < \epsilon;$$

and $$\frac{1}{n+1}\sum_0^n f_1(s_m) \to \int f_1\,dx, \qquad \frac{1}{n+1}\sum_0^n f_2(s_m) \to \int f_2\,dx.$$

Also

$$\varliminf \frac{1}{n+1}\sum_0^n f(s_m) \geqslant \lim \frac{1}{n+1}\sum_0^n f_1(s_m) = \int f_1\,dx > \int f(x)\,dx - \epsilon,$$

$$\varlimsup \frac{1}{n+1}\sum_0^n f(s_m) \leqslant \lim \frac{1}{n+1}\sum_0^n f_2(s_m) = \int f_2\,dx < \int f(x)\,dx + \epsilon;$$

and therefore $$\lim \frac{1}{n+1}\sum_0^n f(s_m) = \int f(x)\,dx,$$

which proves the theorem.

† We use the same symbol for an interval and its length. In what follows an integral without limits shown is over (0, 1).

We can find another criterion for uniform distribution as follows. If

$$f(x) = e^{2k\pi i x} = e(kx),$$

where k is a positive integer, then $\int f(x)\,dx = 0$. It follows that

(5.17.2) $$\sum_0^n e(ks_m) = o(n) \quad (k = 1, 2, 3, \ldots),$$

if (s_n) is uniformly distributed, or, what is the same thing, that

(5.17.3) $$\sum_0^n T(s_m) = o(n),$$

where $T(x)$ is any trigonometrical polynomial without constant term. Thus (5.17.2) or (5.17.3) is a necessary condition for the uniform distribution of (s_n). We now show that the condition is also sufficient.

THEOREM 60. *If* (5.17.2) *is true for every positive integral k, then (s_n) is uniformly distributed.*

First, (5.17.3) is true for every $T(x)$. If

$$\tau(x) = \tfrac{1}{2}a_0 + T(x) = \tfrac{1}{2}a_0 + \sum_{l=1}^{k} (a_l \cos 2l\pi x + b_l \sin 2l\pi x)$$

is any trigonometrical polynomial, then plainly

$$\frac{1}{n+1} \sum_{m=0}^{n} \tau(s_m) \to \tfrac{1}{2}a_0 + \int T(x)\,dx = \int \tau(x)\,dx,$$

and (5.17.1) is true for $f(x) = \tau(x)$.

Next, if $f(x)$ is any real continuous function, there is a τ such that $|f-\tau| < \epsilon$ in $(0, 1)$. If $\tau_1 = \tau-\epsilon$, $\tau_2 = \tau+\epsilon$, then $\tau_1 < f < \tau_2$ and $\int \tau_1\,dx$, $\int \tau_2\,dx$ differ by 2ϵ. It then follows, as in the proof of Theorem 59, that (5.17.1) is true for f. Finally, if $f(x) = I(x)$, then there are continuous functions f_1 and f_2 such that $f_1 \leqslant f \leqslant f_2$ and $\int f_1\,dx$, $\int f_2\,dx$ differ by less than ϵ; and a repetition of the argument shows that (5.17.1) is true also for this f. Hence (s_n) is uniformly distributed.

Perhaps the most interesting case is that in which

$$s_n = n\alpha - [n\alpha] = \{n\alpha\},$$

where α is irrational. If α is a rational p/q, then s_n repeats the cycle of values 0, $1/q$, $2/q, \ldots, (q-1)/q$, in some order, indefinitely. It is therefore natural to expect (s_n) to be uniformly distributed when α is irrational. In this case

$$\sum_0^n e(ks_m) = \sum_0^n e^{2km\pi\alpha i} = \frac{1-e^{2k(n+1)\pi\alpha i}}{1-e^{2k\pi\alpha i}} = O(1) = o(n),$$

for $k = 1, 2, 3, \ldots$. Thus (s_n) is uniformly distributed, and we have

THEOREM 61. *If α is irrational, then the sequence $(\{n\alpha\})$ is uniformly distributed in* $(0, 1)$.

5.18. The uniform distribution of $\{n^2\alpha\}$. There are important generalizations of Theorem 61. In particular, Weyl has shown that $\{P(n)\}$ is uniformly distributed whenever $P(n)$ is a polynomial

$$\alpha_0 n^p + \alpha_1 n^{p-1} + \ldots + \alpha_{p-1} n$$

with at least one irrational coefficient. The proof is a good deal more difficult, and we confine ourselves to a special case, which is sufficient to illustrate Weyl's main idea.

THEOREM 62. *If α is irrational, then the sequence $(\{n^2\alpha\})$ is uniformly distributed in* $(0, 1)$.

We have to prove (5.17.2), with $s_m = \{m^2\alpha\}$, and, since $k\alpha$ is irrational when α is irrational, it is sufficient to prove that

$$S_n = \sum_{m=0}^{n} e^{2m^2\pi\alpha i} = o(n).$$

Now $$|S_n|^2 = \sum_{p=0}^{n} \sum_{q=0}^{n} e^{2(q^2-p^2)\pi\alpha i} = \sum_{p=0}^{n} \sum_{j=-p}^{n-p} e^{2j(j+2p)\pi\alpha i},$$

on writing $p+j$ for q. Inverting the order of summation, we find

$$|S_n|^2 = \sum_{j=0}^{n} e^{2j^2\pi\alpha i} \sum_{p=0}^{n-j} e^{4pj\pi\alpha i} + \sum_{j=-n}^{-1} e^{2j^2\pi\alpha i} \sum_{p=-j}^{n} e^{4pj\pi\alpha i} = T_1 + T_2.$$

Here $$|T_1| \leqslant \sum_{j=0}^{n} \left| \sum_{p=0}^{n-j} e^{4pj\pi\alpha i} \right| = \sum_{j=0}^{n} \left| \frac{1 - e^{4(n-j+1)j\pi\alpha i}}{1 - e^{4j\pi\alpha i}} \right| = \sum_{j=0}^{n} w_j,$$

and w_j satisfies both the inequalities

$$0 \leqslant w_j \leqslant n-j+1 \leqslant n+1, \qquad w_j \leqslant |\operatorname{cosec} 2j\pi\alpha|.$$

Now $|\sin 2j\pi\alpha| \geqslant 2\lambda_j$, where λ_j is the distance of $2j\alpha$ from the nearest integer, i.e. of $\{2j\alpha\}$ from the nearer of 0 and 1. Since the numbers $\{2j\alpha\}$ are uniformly distributed, the number of them with $j \leqslant n$ and $\lambda_j < \eta$, and so lying in one of the intervals $(0, \eta)$ or $(1-\eta, 1)$, is less than $3\eta n$, for sufficiently large n; and then $\omega_j \leqslant (2\eta)^{-1}$ for more than $n+1-3\eta n$ of the j, while $\omega_j \leqslant n+1$ for the remainder. Thus

$$\varlimsup_{n\to\infty} \frac{T_1}{n^2} \leqslant \lim_{n\to\infty} \left\{ \frac{n+1}{2\eta n^2} + \frac{3\eta n(n+1)}{n^2} \right\} = 3\eta,$$

and $T_1 = o(n^2)$. Similarly $T_2 = o(n^2)$; and so $S_n = o(n)$, which proves the theorem.

NOTES ON CHAPTER V

§§ 5.2–3. See § 1.3. The papers of Frobenius and Hölder were published in *JM*, 89 (1880), 262–4, and *MA*, 20 (1882), 535–49; and that of Cesàro in *BSM* (2), 14 (1890), 114–20. Cesàro is concerned primarily with the multiplication of series: see Ch. X.

§ 5.5. The definitions for general k were given independently by Knopp, *Sitzungsberichte d. Berliner Math. Ges.*, 7 (1907), 1–12 [printed in *Archiv d. Math.* (3), 12 (1907)], and by Chapman, *PLMS* (2), 9 (1911), 369–409.

There are general accounts of the theory in the books of Borel, Dienes, Hobson (2, ch. 1), and Knopp, and in the monographs of Andersen, Bohr, and Kogbetliantz. There is also a very clear account of the fundamental theorems in a lecture by Andersen (*Cesàro's Summabilitetsmetode*, Copenhagen, 1919). The monograph of Kogbetliantz is the most complete, but is a summary of results without proofs.

It is sometimes difficult to assign particular theorems to their discoverers, since most of them have been found by a process of gradual generalization; and we do not attempt to do so systematically, though we give the most obvious references.

§ 5.6. The substance of the theorems of this section is Cesàro's. More general theorems of the same character will be found in Knopp, *RP*, 32 (1911), 95–110.

§ 5.7. Theorems 43 and 46, in their general form, are due to Chapman and Knopp. Theorem 45 was proved by Hardy and Littlewood, *PLMS* (2), 11 (1912), 411–78 (462, Theorem 37).

§ 5.8. Knopp, *Grenzwerte von Reihen bei der Annäherung an die Konvergenzgrenze* (Dissertation, Berlin, 1907), proved the implication $(H,k) \rightarrow (C,k)$, and Schnee, *MA*, 67 (1909), 110–25, the converse implication. The proof here is due to Andersen, *MZ*, 28 (1928), 356–9, and is a simplification of one given earlier by Knopp, ibid. 19 (1924), 97–113. See also Knopp, 481.

Theorem 49 is a special case of the theorem that the three hypotheses

$$(a)\ C_n^{(\alpha)}\{C^{(\beta)}(A)\} \to A, \qquad (b)\ C_n^{(\beta)}\{C^{(\alpha)}(A)\} \to A, \qquad (c)\ C_n^{(\alpha+\beta)}(A) \to A$$

are equivalent. This has been proved in various ways by Andersen, Faber, Hausdorff, and Kogbetliantz: references will be found in Andersen's paper. It should be noted that the equivalence of (c) with (a) and (b) lies deeper than that of (a) with (b), the transformations $C^{(\alpha)}C^{(\beta)}$ and $C^{(\beta)}C^{(\alpha)}$ being identical with one another, but not with $C^{(\alpha+\beta)}$.

The identity of $C^{(\alpha)}C^{(\beta)}$ and $C^{(\beta)}C^{(\alpha)}$ is a corollary of Hausdorff's work (Ch. XI), and may also be proved independently. It is easily verified that

$$C_n^{(\alpha)}\{C^{(\beta)}(A)\} = \sum c_{n,p} A_p,$$

where $c_{n,p}$ is 0 for $p > n$ and

$$\frac{\Gamma(n+1)\Gamma(\alpha+1)}{\Gamma(n+\alpha+1)} \frac{\Gamma(p+1)\Gamma(\beta+1)}{\Gamma(p+\beta+1)} \frac{\Gamma(n-p+\alpha)}{\Gamma(n-p+1)\Gamma(\alpha)} {}_3F_2\begin{pmatrix} p+1, & -n+p, & \beta \\ p+\beta+1, & -n+p-\alpha+1 \end{pmatrix}$$

for $p < n$, the argument of the hypergeometric series being 1; and it follows from Bailey, 21, formula (1), that this is symmetrical in α and β.

§ 5.9. Mercer, *PLMS* (2), 5 (1906), 206–24; Schur, *MA*, 74 (1913), 447–58. Schur's proof is also given in Landau, *Ergebnisse*, 43–51.

§ 5.10. Knopp, *MA*, 74 (1913), 459–61; Hardy, *QJM*, 43 (1912), 143–50. Hardy proves a number of extensions of Theorems 52 and 53. The simple proof of Theorem 53 given here is due to Hobson.

Pitt [*PCPS*, 34 (1938), 510–20] and Rogosinski [ibid. 38 (1942), 166–92 and 344–63] have proved much more general theorems by deeper methods. These depend on the use of Fourier and Mellin transforms in the manner of Wiener (Ch. XII).

§ 5.11. Theorem 54 is due to Schur, l.c. under § 5.9.

It follows from the analysis here and Theorem 11 that the (H, k) kernel of (s_n) is included in the (C, k) kernel. Knopp, l.c. under § 3.7, gives a simple example of a real (s_n) whose (H, 2) and (C, 2) kernels are the intervals $(\frac{1}{4}, \frac{3}{4})$ and $(0, 1)$.

Bosanquet, *JLMS*, 21 (1946), 11–15, has shown that $s_n \to \infty$ (H, 2) does not imply $s_n \to \infty$ (C, k) for *any* k, or $s_n \to \infty$ (A), even when $\sum a_n x^n$ is convergent for $|x| < 1$.

Basu, *PLMS* (2), 50 (1948), 447–62, has proved that Theorem 54 remains true for general $k > 1$ and for $-1 < k < 0$, but that the relations are inverted when $0 < k < 1$.

§ 5.12. Theorem 57 is due to Appell, *Archiv d. Math.* 64 (1879), 387–92. The example used to prove Theorem 56 is Landau's (l.c. under § 5.9, 51).

§ 5.13. For the 'harmonic' means see M. Riesz, l.c. under § 4.3.

§§ 5.14–15. It is difficult to give useful references for theorems concerning summable integrals, since they have been often dismissed as 'obvious analogues' of theorems about series. The equivalence theorem was proved first by Landau, *Leipziger Berichte*, 65 (1913), 131–8. Landau's proof is modelled on Schur's of § 5.9.

M. E. Grimshaw, *JLMS*, 9 (1934), 94–102, proves the analogue of Theorem 45. Some further references are given in the notes on Chs. VI and X.

In the text we suppose for simplicity that $a(x)$ is bounded in every finite $(0, X)$. The analysis for Hölder means is valid for all integrable $a(x)$. The same is true for Cesàro means with $k > 0$, but the integrals which occur may sometimes diverge when $k < 0$. Thus $\int (x-t)^k a(t)\,dt$ diverges for $x = n\pi$ when $a(x) = (\sin x)^{-\frac{1}{4}}$ and $-1 < k \leqslant -\frac{3}{4}$. This is unimportant here, since the means of negative order are only interesting in themselves when $a(x)$ tends to a limit.

There is a full discussion of the formula (5.14.3), for $a(x)$ integrable in the more general Denjoy-Perron sense, in Bosanquet, *PLMS* (2), 31 (1930), 144–64.

The $A(x)$ of the text, being the integral of $a(x)$, is absolutely continuous. But we may plainly define $A(x) \to A$(C, k) by (5.14.2) whenever $A(x)$ is integrable, provided that $k > 0$ and we use the first form of the integral. On the other hand, the integrability of $A(x)$ down to 0 does not necessarily imply that of $H^1(x)$: thus $H^1(x) = x^{-1}\left(\log\frac{1}{x}\right)^{-1}$ when $A(x) = x^{-1}\left(\log\frac{1}{x}\right)^{-2}$. We must therefore impose some additional restriction on $A(x)$ for small x. Since we are interested primarily in large x, this is no serious drawback.

§ 5.16. The equivalence of the (R, n, k) and (C, k) means was first proved by M. Riesz, *CR*, 152 (1911), 1651–4: the lines of the proof are indicated rather shortly. There is a complete proof in Hobson (2), 90–8. A more concise version, by Ingham, has not been published; this reduces, when k is an integer, to the proof in the text.

§ 5.17. Theorem 59 was proved independently, at about the same time, by Bohl, Sierpinski, and Weyl: references will be found in Koksma. We follow Weyl, *MA*, 77 (1916), 313–52.

There is an 'elementary' proof of Theorem 61, depending on simple properties of continued fractions, in Hardy and Wright, 378–80.

Weyl proves much more, in particular the uniform distribution of the points

$$\{P_1(n)\},\quad \{P_2(n)\},\quad \ldots,\quad \{P_r(n)\}$$

in r-dimensional space; here $P_1(n),\ldots$ are polynomials linearly independent in the sense that no combination $\lambda_1 P_1+\lambda_2 P_2+\ldots+\lambda_r P_r$, with integral λ, is congruent to a constant (mod 1).

A number of special cases of Weyl's theorems had been stated earlier by Hardy and Littlewood. Thus they state [*Proc. fifth international congress of mathematicians*, Cambridge, 1912, 1, 223–9 (226); *AM*, 37 (1914), 155–91 (164)] that

$$\text{(a)} \qquad \sum_0^n e(m^p\alpha) = o(n)$$

for $p = 1, 2,\ldots$ and irrational α, and that the points $\{n^p\alpha\}$ are uniformly distributed. In a second paper in *AM* (ibid. 193–239) they prove (*a*) for $p = 2$ by a special method, and more precise results for particular types of irrationals. A third paper which was to contain the proofs of their more general assertions was never completed because of the appearance of Weyl's more compact and powerful analysis.

In their first paper in *AM* Hardy and Littlewood prove, by more elementary reasoning, that the points $\{n^p\alpha\}$ are *dense* in $(0, 1)$: this is, of course, a weaker assertion than uniform distribution. Their argument was simplified and extended by Kakeya, *Science reports Tôhoku Univ.* 4 (1915), 105–9.

VI

ARITHMETIC MEANS (2)

6.1. Tauberian theorems for Cesàro summability. We remarked in § 3.8 that there must be a 'limitation theorem' for every method of summation, since no useful method will sum too rapidly divergent series. Thus the limitation theorem for the Cesàro methods is Theorem 46, with $k' = -1$.

There is another limit, of a less obvious kind, to the effectiveness of these methods, and of all that have proved useful. Every method will fail to sum series which diverge too rapidly; and it will also fail to sum *divergent series whose divergence is too slow*. The theorems which embody this principle belong to the class which (for reasons which will appear later) are called 'Tauberian'. They assert that if a series is summable (P), and satisfies some further condition K_P (which will vary with the method P, but will in any case imply a certain slowness of possible divergence), then it is convergent. For the Cesàro methods the most characteristic form of K_P is $a_n = O(n^{-1})$, though this form may be generalized in various ways.

We shall prove the following two theorems.

THEOREM 63. *If* $\sum a_n = A$ (C, k) *for some* k, *and*

$$a_n = O(n^{-1}), \tag{6.1.1}$$

then $\sum a_n$ *is convergent, and indeed summable* (C, $-1+\delta$) *for every positive* δ.

THEOREM 64. *If* a_n *is real,* $\sum a_n = A$ (C, k) *for some* k, *and*

$$na_n > -H, \tag{6.1.2}$$

then $\sum a_n$ *is convergent.*

We can simplify the argument by a few preliminary remarks. First, after Theorem 43, we may suppose k integral, replacing k by $k' = [k]+1$ otherwise. Next, we need only prove the series convergent, since if it is convergent, and satisfies (6.1.1), it is summable (C, $-1+\delta$), by Theorem 45. Finally, we may suppose a_n real, otherwise considering real and imaginary parts separately. Thus it is sufficient to prove Theorem 64, with k integral.

We base the proof on two preliminary theorems of some intrinsic interest. We write $b_n = na_n$, and B_n, B_n^1,... for the sums formed from b_n as A_n, A_n^1,... are from a_n.

THEOREM 65. *If* $\sum a_n$ *is summable* $(\mathrm{C}, r+1)$, *where* $r > -1$, *then a necessary and sufficient condition that it should be summable* (C, r) *is that* $B_n^r = o(n^{r+1})$.

THEOREM 66. *A necessary and sufficient condition that* $\sum a_n$ *should be summable* $(\mathrm{C}, r+1)$, *where* $r+1 > -1$, *is that*

$$\sum \binom{n+r+1}{r+1}^{-1} \frac{B_n^r}{n} = \sum \frac{(n-1)!}{(r+2)(r+3)\dots(n+r+1)} B_n^r \tag{6.1.3}$$

should be convergent; or, what is the same thing, that $\sum n^{-r-2} B_n^r$ *should be convergent.*

It is easily verified that

$$(n+r+1)\binom{\nu+r}{r} - (r+1)\binom{\nu+r+1}{r+1} = (n-\nu)\binom{\nu+r}{r},$$

$$n\binom{\nu+r+1}{r+1} - (n+r+1)\binom{\nu+r}{r+1} = (n-\nu)\binom{\nu+r}{r};$$

and hence (comparing the coefficients of $a_{n-\nu}$ in (5.4.5)) that

$$(n+r+1)A_n^r - (r+1)A_n^{r+1} = B_n^r, \tag{6.1.4}$$

$$nA_n^{r+1} - (n+r+1)A_{n-1}^{r+1} = B_n^r. \tag{6.1.5}$$

From (6.1.4) and (6.1.5) we deduce

$$\binom{n+r}{r}^{-1} A_n^r - \binom{n+r+1}{r+1}^{-1} A_n^{r+1} = \binom{n+r+1}{r+1}^{-1} \frac{B_n^r}{r+1}, \tag{6.1.6}$$

$$\binom{n+r+1}{r+1}^{-1} A_n^{r+1} - \binom{n+r}{r+1}^{-1} A_{n-1}^{r+1} = \binom{n+r+1}{r+1}^{-1} \frac{B_n^r}{n};$$

and addition of the last equation for $n = 1, 2, \dots, N$ gives

$$\binom{N+r+1}{r+1}^{-1} A_N^{r+1} = a_0 + \sum_1^N \binom{n+r+1}{r+1}^{-1} \frac{B_n^r}{n}. \tag{6.1.7}$$

Theorem 65 is a corollary of (6.1.6), and Theorem 66 of (6.1.7). The two forms of Theorem 66 are equivalent by Stirling's theorem. If r is an integer, then the series (6.1.3) may be written in the alternative form

$$\sum \frac{(r+1)!}{n(n+1)\dots(n+r+1)} B_n^r.$$

We can now prove Theorem 64: we may suppose k an integer $r+1$, and $H = 1$. If $B_n^r \neq o(n^{r+1})$, then there is a positive C such that one or other of the inequalities

$$B_n^r > Cn^{r+1}, \tag{6.1.8}$$ $$B_n^r < -Cn^{r+1} \tag{6.1.9}$$

is true for an infinity of n. Let us suppose, for example, that (6.1.8) is true for an infinity of values N of n.

If $\eta > 1$ and $N \leqslant n \leqslant \eta N$, then

$$(6.1.10)\quad B_n^r - B_N^r = \sum_{\nu=1}^{N} \left\{\binom{n-\nu+r}{r} - \binom{N-\nu+r}{r}\right\} b_\nu + \sum_{\nu=N+1}^{n} \binom{n-\nu+r}{r} b_\nu.$$

Hence, since the coefficients are positive, and $b_\nu > -1$,

$$B_n^r - B_N^r > -\sum_{\nu=1}^{N} \left\{\binom{n-\nu+r}{r} - \binom{N-\nu+r}{r}\right\} - \sum_{\nu=N+1}^{n} \binom{n-\nu+r}{r}.$$

Here the right-hand side is what stands in (6.1.10) when $b_0 = 0$, $b_\nu = -1$ for $\nu > 0$, in which case

$$\sum B_n^r x^n = (1-x)^{-r-1} \sum b_n x^n = -x(1-x)^{-r-2}, \qquad B_n^r = -\binom{n+r}{r+1};$$

and hence
$$B_n^r - B_N^r > -\binom{n+r}{r+1} + \binom{N+r}{r+1}.$$

Also
$$\binom{n+r}{r+1} \sim \frac{n^{r+1}}{(r+1)!}, \qquad \binom{N+r}{r+1} \sim \frac{N^{r+1}}{(r+1)!},$$

and therefore

$$B_n^r - B_N^r > -\frac{1}{(r+1)!}\{(1+\epsilon)\eta^{r+1} - (1-\epsilon)\} N^{r+1}$$

for any positive ϵ, any $\eta > 1$, $N \leqslant n \leqslant \eta N$, and sufficiently large N.

We can choose ϵ and η so that $B_n^r - B_N^r > -\frac{1}{2}CN^{r+1}$, and it then follows from (6.1.8), with $n = N$, that $B_n^r > \frac{1}{2}CN^{r+1}$ for $N \leqslant n \leqslant \eta N$, and so

$$\sum_N^{\eta N} \frac{B_n^r}{n^{r+2}} > \tfrac{1}{2}CN^{r+1} \sum_N^{\eta N} \frac{1}{n^{r+2}} > \tfrac{1}{2}CN^{r+1} \frac{(\eta-1)N}{(\eta N)^{r+2}} = \frac{C(\eta-1)}{2\eta^{r+2}}$$

for sufficiently large N. But if this is true for an infinity of N, then the series (6.1.3) is divergent, and $\sum a_n$ is not summable $(C, r+1)$.

It follows that (6.1.8) cannot be true for an infinity of n, and a similar argument† shows that (6.1.9) cannot. Hence $B_n^r = o(n^{r+1})$; and therefore, by Theorem 65, $\sum a_n$ is summable (C, r). Repeating the argument $r+1$ times, we see that $\sum a_n$ is convergent.

It will be observed that Theorem 63 goes farther than Theorem 64, in that it asserts summability of the series for negative k. No such extension of Theorem 64 is possible, since the conditions are satisfied by any series of positive terms, and, after Theorem 46, $\sum a_n$ cannot be summable $(C, -l)$ unless $a_n = o(n^{-l})$.

There are generalizations of these theorems for Riesz's typical means of § 4.16. We shall not consider these here, except for one theorem which

† Using a range $(\zeta N, N)$, where $\zeta < 1$, of values of n.

we shall have to use later. This is a generalization of the case $k = 1$ of Theorem 63.

THEOREM 67. *If* $0 \leqslant \lambda_0 < \lambda_1 < \ldots$, $\lambda_n \to \infty$,

$$a_n = O\left(\frac{\lambda_n - \lambda_{n-1}}{\lambda_n}\right) \quad (n > 0), \tag{6.1.11}$$

and

$$\frac{1}{x}\int_0^x A(u)\,du = \frac{1}{x}\int_0^x \left(\sum_{\lambda_n \leqslant u} a_n\right) du \to s, \dagger \tag{6.1.12}$$

then $\sum a_n$ *converges to* s.

We may suppose that $s = 0$ and $|a_n| < (\lambda_n - \lambda_{n-1})/\lambda_n$ for $n > 0$. If $t > x$ and $\lambda_m \leqslant x < \lambda_{m+1}$, $\lambda_{m+r} \leqslant t < \lambda_{m+r+1}$, then

$$|A(t) - A(x)| = |a_{m+1} + a_{m+2} + \ldots + a_{m+r}| \tag{6.1.13}$$

$$\leqslant \frac{\lambda_{m+1} - \lambda_m}{\lambda_m} + \ldots + \frac{\lambda_{m+r} - \lambda_{m+r-1}}{\lambda_{m+r}} \leqslant \frac{\lambda_{m+r} - \lambda_m}{\lambda_m} \leqslant \frac{t - \lambda_m}{\lambda_m}.$$

If $A(x)$ does not tend to 0, then there is a $C > 0$ such that one or other of $A(x) > C$, $A(x) < -C$ is true for a sequence of values X of x tending to infinity. If, for example, $A(X) > C$ and $\lambda_M \leqslant X < \lambda_{M+1}$, so that $A(X) = A(\lambda_M)$, then, by (6.1.13),

$$A(t) > C - (t - \lambda_M)/\lambda_M > \tfrac{1}{2}C$$

for $\lambda_M < t < (1 + \frac{1}{2}C)\lambda_M$, and hence

$$\int_{\lambda_M}^{(1+\frac{1}{2}C)\lambda_M} A(t)\,dt > \tfrac{1}{4}C^2\lambda_M,$$

in contradiction to (6.1.12). Similarly $A(X) < -C$ leads to a contradiction, and so $A(x) \to 0$.

We cannot replace (6.1.11) by $a_n > -H(\lambda_n - \lambda_{n-1})/\lambda_n$ without some further restriction either on λ_n or on a_n.‡

6.2. Slowly oscillating and slowly decreasing functions. A function $f(x)$, defined for $x > 0$, is said to be *slowly oscillating* if

$$f(y) - f(x) \to 0 \tag{6.2.1}$$

whenever

$$x \to \infty, \qquad y > x, \qquad y/x \to 1; \tag{6.2.2}$$

and to be *slowly decreasing* if it is real and

$$\varliminf \{f(y) - f(x)\} \geqslant 0 \tag{6.2.3}$$

† $A(u)$ here is equivalent to the $A_\lambda(u)$ of § 4.16.

‡ See the note on this section at the end of the chapter, and that on § 7.7.

in the same circumstances. If $f(x)$ is differentiable, and $f'(x) = O(x^{-1})$, then

$$f(y)-f(x) = \int_x^y f'(t)\,dt = O\left(\frac{y-x}{x}\right) \to 0$$

under the conditions (6.2.2), so that $f(x)$ is slowly oscillating. Thus $x^{ai} = e^{ai\log x}$ is slowly oscillating. Similarly, if $f(x)$ is real and $f'(x) > -Hx^{-1}$, then $f(x)$ is slowly decreasing: thus $x+\cos x+\cos(a\log x)$ is slowly decreasing.

We shall say that a sequence s_n is slowly oscillating, or slowly decreasing, if $s(x) = s_{[x]}$ is slowly oscillating or slowly decreasing. If $s_n = a_0+a_1+...+a_n$, then $s(x)$ is the sum-function of $\sum a_n$. It is easily verified that s_n is slowly oscillating when $a_n = O(n^{-1})$, slowly decreasing when $a_n > -Hn^{-1}$.

If $f(x)$ is slowly oscillating, then $|f(y)-f(x)| < \epsilon$ when $y > x \geqslant X(\epsilon)$ and $(y-x)/x \leqslant \kappa(\epsilon)$. If it is slowly decreasing, then $f(y)-f(x) > -\epsilon$ under similar conditions.

There is one simple corollary which we shall require in Ch. VII. If $f(x)$ is slowly decreasing, and $q > 0$, $p > q$ are fixed, then there are an H and an X such that

(6.2.4) $$f(px)-f(qx) > -H$$

for $x \geqslant X$. For there are a U and a κ such that

$$f(t)-f(u) > -1 \quad (qu \geqslant U,\ 1 < t/u \leqslant \kappa).$$

If r is the integer for which $\kappa^{r-1} < p/q \leqslant \kappa^r$, and

$$x_0 = qx, \quad x_1 = \kappa qx, \ldots, \quad x_{r-1} = \kappa^{r-1}qx, \quad x_r = px,$$

then we may take $t = x_{s+1}$, $u = x_s$ for $s = 0, 1, \ldots, r-1$, and

$$f(px)-f(qx) = \sum_{s=0}^{r-1}\{f(x_{s+1})-f(x_s)\} > -r,$$

so that (6.2.4) is satisfied with $X = U/q$, $H = r$. If also $f(x)$ is bounded in every finite interval $(0, X)$, then (6.2.4) is satisfied, with an appropriate H, for $x \geqslant 0$.

There are important generalizations of Theorems 63 and 64 in which the condition on a_n is replaced by the more general condition that s_n is slowly oscillating or slowly decreasing. These will be included in the more difficult theorems proved in Ch. VII, but we illustrate the ideas here by proving the simplest theorem of this kind.

THEOREM 68. *If $\sum a_n$ is summable* (C, 1), *and s_n is slowly decreasing, then $\sum a_n$ is convergent.*

We are given that $s_n \to s$ (C, 1) and that

$$\underline{\lim}\,(s_{pn}-s_n) \geqslant 0$$

when $n \to \infty$, $p > 1$, and $p \to 1$; and it is sufficient, after Theorem 65, to prove that

$$u_n = a_1+2a_2+\ldots+na_n = o(n). \tag{6.2.5}$$

Let us suppose, for example, that

$$u_n > Cn \tag{6.2.6}$$

for a positive C and an infinity of n, in contradiction to (6.2.5). Given any positive η, we can choose N and $p > 1$ so that $s_\nu - s_n > -\eta$ for $n \geqslant N$ and $n \leqslant \nu \leqslant pn$; and we may suppose that $p\eta < \frac{1}{2}C$. Then

$$u_n = (n+1)s_n - s_0 - s_1 - \ldots - s_n,$$

$$u_\nu - u_n = (n+1)(s_\nu - s_n)+(s_\nu - s_{n+1})+\ldots+(s_\nu - s_{\nu-1}) \geqslant -\nu\eta \geqslant -p\eta n,$$

and
$$u_\nu = u_n + u_\nu - u_n > Cn - p\eta n > \tfrac{1}{2}Cn$$

for any $n \geqslant N$ satisfying (6.2.6) and $n \leqslant \nu \leqslant pn$. Thus

$$\sum_{\nu=n}^{[pn]} \frac{u_\nu}{\nu(\nu+1)} > \tfrac{1}{2}Cn \sum_{\nu=n}^{[pn]} \frac{1}{\nu(\nu+1)} = \tfrac{1}{2}Cn\left(\frac{1}{n} - \frac{1}{[pn]+1}\right) > \tfrac{1}{2}C\left(1-\frac{1}{p}\right).$$

Hence the series $\sum \dfrac{u_n}{n(n+1)}$ is not convergent, and so, by Theorem 66, with $r = 0$, $\sum a_n$ is not summable (C, 1).

Similarly, we can show that the hypothesis $u_n < -Cn$, for an infinity of n, leads to a contradiction. Thus $u_n = o(n)$, and the theorem follows.

The corresponding theorem for functions of a continuous variable is

if $f(t) \to l$ (C, 1) and $f(t)$ is slowly decreasing, then $f(t) \to l$:†

the proof is left to the reader.

6.3. Another Tauberian condition. There are conditions of other types which enable us to infer convergence from summability. As an example, we prove

THEOREM 69. *If $\sum a_n$ is summable (C, l) for some l, $p \geqslant 1$, and $\sum n^{p-1}|a_n|^p < \infty$, then $\sum a_n$ is convergent, and indeed summable (C, k) for $k > -(p-1)/p$.*

The result is trivial when $p = 1$, and we may suppose $p > 1$. It is sufficient, after Theorem 65, to prove that

$$B_n^k = \sum_{\nu=0}^{n} \binom{n-\nu+k}{k} \nu a_\nu = o(n^{k+1}) \quad \left(k > -1+\frac{1}{p}\right). \tag{6.3.1}$$

† This is what, in the notation laid down in Ch. VII, we should call Theorem 68 *a*.

Now $$B_n^k = O\left\{\sum_{\nu=0}^{n}(n-\nu+1)^k\nu|a_\nu|\right\} = O\left(\sum_{\nu=0}^{N}+\sum_{\nu=N+1}^{n}\right) = S_1+S_2,$$
say. Here
$$|S_2| \leqslant \sum_{\nu=N+1}^{n}\nu^{(p-1)/p}|a_\nu|.\nu^{1/p}(n-\nu+1)^k$$
$$\leqslant \left(\sum_{\nu=N+1}^{n}\nu^{p-1}|a_\nu|^p\right)^{1/p}\left\{\sum_{\nu=N+1}^{n}\nu^{1/(p-1)}(n-\nu+1)^{kp/(p-1)}\right\}^{(p-1)/p},$$
by Hölder's inequality. The second factor is $O(n^q)$, where
$$q = \left(\frac{kp}{p-1}+\frac{1}{p-1}+1\right)\frac{p-1}{p} = k+1,$$
and the first is less than ϵ for $N \geqslant N_0(\epsilon)$. Hence $|S_2| < C\epsilon n^{k+1}$, where C is independent of n, for $N \geqslant N_0(\epsilon)$; and S_1 is plainly $o(n^{k+1})$ when N is fixed. This proves (6.3.1), and therefore the theorem.

6.4. Convexity theorems. If $\sum a_n$ is summable (C, k) then, by Theorem 43, it is summable (C, k') for any $k' > k$; and if it is bounded (C, k) then it is bounded (C, k'). But boundedness (C, k) does not imply summability (C, k'), for any k'. There is, however, a slightly more subtle theorem.

THEOREM 70. *If $\sum a_n$ is bounded* (C, k_1), *and summable* (C, k_2), *where $k_2 > k_1 > -1$, then it is summable* (C, k) *for $k_1 < k < k_2$.*

We prove this here only for integral k_1, k_2, k, when $k_2 = k_1+l$, $k = k_1+m$, l and m being integers and $0 < m < l$. It is sufficient to prove the theorem when $l = 2$, $m = 1$. For suppose the theorem proved in this case, and also for general l, m with $l = 2, 3,..., L-1$; and consider the case $l = L$. Then $\sum a_n$, being bounded (C, k_1), is bounded (C, k_1+L-2), and therefore (by hypothesis) summable (C, k_1+L-1); and hence (again by hypothesis) it is summable (C, k_1+m) for $0 < m < L$.

We may also suppose that the sum (C, k_2) is 0; and we have therefore to prove that $A_n^k = O(n^k)$ and $A_n^{k+2} = o(n^{k+2})$ imply $A_n^{k+1} = o(n^{k+1})$; or, writing B_n for A_n^k, that $B_n = O(n^k)$ and $B_n^2 = o(n^{k+2})$ imply $B_n^1 = o(n^{k+1})$.

Suppose that $0 < \vartheta < 1$ and $N = [\vartheta n]$. Then
$$B_n^2-B_N^2 = B_{N+1}^1+B_{N+2}^1+...+B_n^1 = (n-N)B_n^1-\sum_{\nu=N+1}^{n}(B_n^1-B_\nu^1)$$
$$= (n-N)B_n^1-\{B_{N+2}+2B_{N+3}+...+(n-N-1)B_n\},$$
$$B_n^1 = \frac{B_n^2-B_N^2}{n-N}+\frac{B_{N+2}+2B_{N+3}+...+(n-N-1)B_n}{n-N} = P+Q,$$
$$Q = O\left\{\frac{1}{(1-\vartheta)n}.n^k\sum_{\nu\leqslant(1-\vartheta)n}\nu\right\} = O\{(1-\theta)n^{k+1}\},$$

uniformly in ϑ, and

$$P = o(n^{-1}.n^{k+2}) = o(n^{k+1})$$

when ϑ is fixed. Hence (taking ϑ near 1) we deduce that $B_n^1 = o(n^{k+1})$.

Theorem 45 is the form assumed by Theorem 70 when $k_1 = -1$, $k_2 = 0$. In this special case we have proved the theorem for non-integral k.

6.5. Convergence factors. A familiar theorem of Abel and Dirichlet, included in Theorem 8 of § 3.5, states that if (i) $\sum a_n$ is convergent or bounded, (ii) f_n decreases steadily to 0 when $n \to \infty$, or, more generally, $f_n \to 0$ and $\sum |\Delta f_n| < \infty$, then $\sum a_n f_n$ is convergent. There are many important generalizations of this theorem for summable series.

These generalizations are of two types. In the first, we impose on f_n only the natural extensions of condition (ii), and infer the summability (C, k) of $\sum a_n f_n$ from that of $\sum a_n$. In the second, we impose stronger conditions on f_n, and infer that $\sum a_n f_n$ is summable $(C, k-s)$ for some positive s: thus a typical case would be that in which $f_n = (n+1)^{-s}$. Both types of theorem present considerable difficulties when the parameters are unrestricted, and we shall confine ourselves here to integral k and s, for which the proofs of the main theorems are comparatively simple.

The principal theorem of the first type is

THEOREM 71. *If* (i) $\sum a_n$ *is summable, or bounded,* (C, k), *where* k *is an integer*; (ii) $f_n \to 0$; *and* (iii)

$$\sum (n+1)^k |\Delta^{k+1} f_n| < \infty; \tag{6.5.1}$$

then $\sum a_n f_n$ *is summable* (C, k), *and*

$$\sum a_n f_n = \sum A_n^k \Delta^{k+1} f_n, \tag{6.5.2}$$

the last series being absolutely convergent.

We require two lemmas.

THEOREM 72. *If* f_n *satisfies the conditions of Theorem* 71, *then*

$$(n+1)^l \Delta^l f_n \to 0 \quad (l = 0, 1, ..., k), \tag{6.5.3}$$

$$\sum (n+1)^l |\Delta^{l+1} f_n| < \infty \quad (l = 0, 1, ..., k). \tag{6.5.4}$$

We can write (6.5.3) and (6.5.4) in the equivalent forms

$$\binom{n+l}{l} \Delta^l f_n \to 0 \quad (l = 0, 1, ..., k), \tag{6.5.5}$$

$$\sum \binom{n+l}{l} |\Delta^{l+1} f_n| < \infty \quad (l = 0, 1, ..., k). \tag{6.5.6}$$

The conditions given are (6.5.5) for $l = 0$ and (6.5.6) for $l = k$.

Since $f_n \to 0$, $\Delta^k f_n \to 0$, and so

$$\Delta^k f_n = \sum_n^\infty \Delta^{k+1} f_\nu,$$

$$(n+1)^k |\Delta^k f_n| \leqslant (n+1)^k \sum_n^\infty |\Delta^{k+1} f_\nu| \leqslant \sum_n^\infty (\nu+1)^k |\Delta^{k+1} f_\nu| \to 0,$$

by (6.5.1). This is (6.5.3), or (6.5.5), for $l = k$.

Next,

$$\sum_{n=0}^\infty \binom{n+k-1}{k-1} |\Delta^k f_n| \leqslant \sum_{n=0}^\infty \binom{n+k-1}{k-1} \sum_{\nu=n}^\infty |\Delta^{k+1} f_\nu|$$

$$= \sum_{\nu=0}^\infty |\Delta^{k+1} f_\nu| \sum_{n=0}^{\nu} \binom{n+k-1}{k-1} = \sum_{\nu=0}^\infty \binom{\nu+k}{k} |\Delta^{k+1} f_\nu| < \infty.$$

It follows that (6.5.6) is true for $l = k-1$; and therefore, by the argument of the preceding paragraph, that (6.5.5) is also true for $l = k-1$. Repeating the argument, we conclude that both assertions are true generally.

THEOREM 73. *If $b_n = a_n f_n$ then*

$$B_n^k = \sum_{j=0}^n A_j^k \sum_{i=0}^k \binom{k+1}{i} \binom{n-j+k-i}{k-i} \Delta^{k+1-i} f_{j+i}, \tag{6.5.7}$$

with the convention that any f_m occurring in $\Delta^{k+1-i} f_{j+i}$ is to be replaced by 0 when $m > n$.

If $U_j = u_0 + u_1 + \ldots + u_j$, U_j^1, U_j^2,... are defined as usual, and v_j with $j > n$ are treated as 0, then

$$\sum_0^n u_j v_j = \sum_0^n U_j \Delta v_j = \sum_0^n U_j^1 \Delta^2 v_j = \ldots = \sum_0^n U_j^k \Delta^{k+1} v_j;$$

and hence

$$B_n^k = \sum_0^n \binom{n-j+k}{k} a_j f_j = \sum_{j=0}^n A_j^k \Delta^{k+1} \left\{ \binom{n-j+k}{k} f_j \right\}.$$

But
$$\Delta^{k+1}(c_j f_j) = \sum_{i=0}^{k+1} \binom{k+1}{i} \Delta^i c_j \Delta^{k+1-i} f_{j+i}.$$

Since here

$$\Delta^i c_j = \Delta^i \binom{n-j+k}{k} = \binom{n-j+k-i}{k-i} \quad (0 \leqslant i \leqslant k),$$

and $\Delta^{k+1} c_j = 0$, we obtain (6.5.7).

Passing to the proof of Theorem 71, we divide (6.5.7) by $\binom{n+k}{k}$ and make $n \to \infty$. First, we may discard the convention. For it affects only

the terms in which $j \geqslant n-k$; the number of such terms is bounded; and each involves an A_j^k which is $O(n^k)$, an f_m which is $o(1)$, and a bounded numerical coefficient; so that their aggregate is $o(n^k)$. We have therefore to find the limit of

$$\binom{n+k}{k}^{-1} \sum_{j=0}^{n} A_j^k \sum_{i=0}^{k} \binom{k+1}{i}\binom{n-j+k-i}{k-i}\Delta^{k+1-i}f_{j+i} = \sum_{i=0}^{k} S_{n,i},$$

where $S_{n,i}$ contains the terms involving a given i.

If $i > 0$ then

$$S_{n,i} = O\left\{(n+1)^{-k}\sum_{j=0}^{n}(j+1)^k\binom{n-j+k-i}{k-i}|\Delta^{k+1-i}f_{j+i}|\right\} = O\Big(\sum_{j=0}^{n} u_{n,j}\Big),$$

where

$$|u_{n,j}| \leqslant H(n+1)^{-k}(j+1)^k(n+1)^{k-i}|\Delta^{k+1-i}f_{j+i}|$$
$$= H\left(\frac{j+1}{n+1}\right)^i (j+1)^{k-i}|\Delta^{k+1-i}f_{j+i}| \leqslant H(j+1)^{k-i}|\Delta^{k+1-i}f_{j+i}|$$

and H is independent of n. Also $\sum_{j=0}^{\infty}(j+1)^{k-i}|\Delta^{k+1-i}f_{j+i}| < \infty$, by Theorem 72. Hence $\sum u_{n,j}$ is majorized by a convergent series with terms independent of n; and $u_{n,j} \to 0$, for any fixed j, when $n \to \infty$. It follows that $S_{n,i} = o(1)$ for $i = 1, 2,..., k$.

It remains to find the limit of

$$S_{n,0} = \binom{n+k}{k}^{-1}\sum_{j=0}^{n} A_j^k \binom{n-j+k}{k}\Delta^{k+1}f_j.$$

This is majorized by $\qquad \sum |A_j^k||\Delta^{k+1}f_j|;$

and $\qquad \binom{n-j+k}{k}\Big/\binom{n+k}{k} \to 1,$

for every j, when $n \to \infty$. Hence

$$S_{n,0} \to \sum A_j^k \Delta^{k+1}f_j,$$

and this completes the proof of Theorem 71.

We may modify Theorem 71 by supposing (i′) that $\sum a_n$ is summable (C, k) and (ii′) that f_n is bounded. The last condition, with (iii), ensures that f_n tends to a limit f, not necessarily 0. The conclusion then follows from Theorem 71 on replacing f_n by $g_n + f$.

Theorem 71, and the modified theorem, may also be deduced from Theorems 1 and 3, and it may be shown that the conditions are also necessary, in the sense that, if they are not satisfied, there are series

$\sum a_n$ which are bounded (C, k), or summable (C, k), while $\sum a_n f_n$ is not summable (C, k).

If $f_n = (n+c)^{-s}$, where $c > 0$, $s > 0$, then $\Delta^{k+1} f_n = O(n^{-s-k-1})$; and if

(6.5.8) $$f_n = c_0 + \frac{c_1}{n+1} + \ldots + \frac{c_{k+1}}{(n+1)^{k+1}} + O\left\{\frac{1}{(n+1)^{k+2}}\right\},$$

then $\Delta^{k+1} f_n = O(n^{-k-2})$. In either of these cases condition (iii) of Theorem 71 is satisfied. We thus obtain

THEOREM 74. *If $\sum a_n$ is summable* (C, k), *then $\sum (n+c)^{-s} a_n$ is summable* (C, k).

THEOREM 75. *If $\sum a_n$ is summable* (C, k), *and f_n is of the form* (6.5.8), *then $\sum a_n f_n$ is summable* (C, k).

In particular

(6.5.9) $$f_n = \frac{(n+\alpha_1)(n+\alpha_2)\ldots(n+\alpha_l)}{(n+\beta_1)(n+\beta_2)\ldots(n+\beta_l)},$$

where the α and β are positive, is of the form (6.5.8), and so is f_n^{-1}; thus the summability of either of

$$\sum \frac{a_n}{(n+\alpha_1)\ldots(n+\alpha_l)}, \qquad \sum \frac{a_n}{(n+\beta_1)\ldots(n+\beta_l)},$$

for any k, involves that of the other. We shall use this case of Theorem 75 in the next section.

6.6. The factor $(n+1)^{-s}$.

The principal theorem of the second type is

THEOREM 76. *If $\sum a_n$ is summable* (C, k), *and*

$$0 < s < k+1,$$

then the series

(6.6.1) $$\sum \binom{n+s}{s}^{-1} a_n, \qquad \sum \frac{a_n}{(n+1)^s}$$

are summable (C, $k-s$).

We suppose k and s integers, all proofs for non-integral values being a good deal more troublesome. Some preliminary remarks are required.

(1) If either of the series $a_0+a_1+a_2+\ldots$ and $0+a_0+a_1+\ldots$ (i.e. $0+b_1+b_2+\ldots$, where $a_n = b_{n+1}$) is summable (C, k), then so is the other, by Theorem 47. It is therefore indifferent whether we state Theorem 76 in terms of $a_0+a_1+\ldots$ and the series (6.6.1), or in terms of $a_1+a_2+\ldots$ and the series

(6.6.2) $$\sum \binom{n-1+s}{s}^{-1} a_n, \qquad \sum \frac{a_n}{n^s}$$

summed over 1 to ∞. It is plain that, if we do this, we may suppose any of the α or β of Theorem 75 to be 0.

(2) Next, $$n^{-s}\binom{n-1+s}{s} = \frac{1}{s!}\frac{(n+1)(n+2)..(n+s-1)}{n^{s-1}}$$
is of the form (6.5.9), with $n-1$ for n and $l = s-1$. Thus the summability (C, r) of either of (6.6.2), for integral r, implies that of the other.

We consider the second series (6.6.2), with $s = 1$. If $\sum a_n$ is summable (C, k), then $\sum n^{-1}a_n$ is summable (C, k), by Theorem 71. In order that it should be summable $(C, k-1)$, it is necessary and sufficient, by Theorem 65, that $Q_n^{k-1} = o(n^k)$, where Q_n^{k-1} is formed from $q_n = n(n^{-1}a_n) = a_n$; i.e. that $A_n^{k-1} = o(n^k)$. But this is true because $A_n^{k-1} = A_n^k - A_{n-1}^k$ and $n^{-k}A_n^k$ tends to a limit.

Thus the series $$\sum n^{-1}a_n$$
is summable $(C, k-1)$. It follows, by Theorem 75, that $\sum (n+\alpha)^{-1}a_n$, where $\alpha \geqslant 0$, is summable $(C, k-1)$. Hence, repeating the argument s times, the series (6.6.2) or (6.6.1) are summable $(C, k-s)$.

It follows from Theorems 71 and 76 that, if $\sum a_n$ is summable (C, k),

(6.6.3)
$$\sum \frac{a_n}{n+1} = \sum A_n^k \Delta^{k+1}\frac{1}{n+1} = (k+1)!\sum \frac{A_n^k}{(n+1)(n+2)...(n+k+2)},$$
the first series being summable $(C, k-1)$ and the last absolutely convergent; there is a similar formula for the sum of the first series (6.6.1).

6.7. Another condition for summability. We saw in § 6.6 (Theorem 76) that the summability (C, k) of $\sum a_n$ implies the summability $(C, k-1)$ of $\sum (n+1)^{-1}a_n$. The converse is false: for the last series is (absolutely) convergent whenever $a_n = O(n^{-1})$, and then, after Theorem 63, $\sum a_n$ cannot be summable (C, k), for any k, unless it is convergent. There is, however, a more subtle connexion between the two series.

THEOREM 77. *If k is integral then, in order that $\sum a_n$ should be summable (C, k), it is necessary and sufficient that there should be a solution b_n of the equations*

(6.7.1) $$a_n = (n+1)(b_n - b_{n+1}) \quad (n = 0, 1, 2,...)$$
such that $\sum b_n$ is summable $(C, k-1)$. In these circumstances

(6.7.2) $$b_n = \frac{a_n}{n+1} + \frac{a_{n+1}}{n+2} + ... \quad (C, k-1),$$

and the sums of the two series are the same.

It is plain, first, that we can solve (6.7.1) by recurrence; that if b_n is any solution then the general solution is $\mathrm{b}_n = b_n - h$, where h is an arbitrary constant; and that the conditions of the theorem can be satisfied by one solution at most.

(i) We begin by proving that, if b_n is any solution, then

(6.7.3) $$A_n^k = (k+1)B_n^k-(n+1)B_{n+1}^{k-1} \quad (k \geqslant 0).$$

Here B_n^{-1} means b_n. First, if $k = 0$, then

$$A_n^0 = A_n = b_0-b_1+2(b_1-b_2)+...+(n+1)(b_n-b_{n+1}) = B_n-(n+1)b_{n+1},$$

which is (6.7.3). Next, assuming (6.7.3) for a given k, we have

$$\begin{aligned} A_n^{k+1} &= (k+1)(B_0^k+B_1^k+...+B_n^k)-B_1^{k-1}-2B_2^{k-1}-...-(n+1)B_{n+1}^{k-1} \\ &= (k+1)B_n^{k+1}-(n+1)B_{n+1}^k+(n+1)B_0^{k-1}+nB_1^{k-1}+...+B_n^{k-1} \\ &= (k+2)B_n^{k+1}-(n+1)B_{n+1}^k, \end{aligned}$$

which is (6.7.3) with $k+1$ for k.

(ii) Next, we prove that *if* $B = \sum b_n$ *is summable* $(C, k-1)$, *then* $A = \sum a_n$ *is summable* (C, k), *and* $A = B$. First, if $k = 0$, then $B_n \to B$ and $(n+1)b_{n+1} \to 0$, so that $A_n \to B$. Secondly, if $k > 0$, then

$$B_n^{k-1} = \binom{n+k-1}{k-1}B+o(n^{k-1})$$

and so, by summation,

$$B_n^k = \binom{n+k}{k}B+o(n^k).$$

Hence, by (6.7.3),

$$A_n^k = \left\{(k+1)\binom{n+k}{k}-(n+1)\binom{n+k}{k-1}\right\}B+o(n^k) = \binom{n+k}{k}B+o(n^k),$$

and A is summable (C, k) to sum B.

(iii) Thirdly, we prove that *if* $\sum a_n$ *is summable* (C, k), *then*

(6.7.4) $$B_n^{k-1} = \binom{n+k}{k}h+\binom{n+k-1}{k-1}A+o(n^{k-1}) \quad (k > 0),$$

(6.7.5) $$b_n = h+o(n^{-1}) \quad (k = 0),$$

h being a constant. We may suppose without loss of generality that $A = 0$.†

We have

(6.7.6)

$$(n+k+2)B_n^k-(n+1)B_{n+1}^k = (k+1)B_n^k-(n+1)B_{n+1}^{k-1} = A_n^k = o(n^k),$$

† If b_n is a solution of (6.7.1), $a_0' = a_0 - A$ and $a_n' = a_n$ for $n > 0$, then the b_n' defined by $b_0' = b_0 - A$, $b_n' = b_n$ for $n > 0$, is a solution of $a_n' = (n+1)(b_n' - b_{n+1}')$. The effect of diminishing a_0 and b_0 by A is to diminish A_n^{k-1} and B_n^{k-1} by $\binom{n+k-1}{k-1}A$.

by (6.7.3) and our hypothesis. If

(6.7.7) $$(n+1)(n+2)...(n+k+1)\phi_n = B_n^k,$$

then (6.7.6) gives

$$(n+1)(n+2)...(n+k+2)(\phi_n-\phi_{n+1}) = A_n^k = o(n^k),$$

and so $\phi_n-\phi_{n+1} = o(n^{-2})$. Hence ϕ_n tends to a limit ϕ, and

(6.7.8) $$\phi_n = \phi+\sum_n^\infty \frac{A_\nu^k}{(\nu+1)(\nu+2)...(\nu+k+2)} = \phi+o\left(\frac{1}{n}\right),$$

(6.7.9) $$B_n^k = (n+1)(n+2)...(n+k+1)\phi+o(n^k).$$

Finally, from (6.7.6) and (6.7.9) it follows that

$$B_{n+1}^{k-1} = \frac{k+1}{n+1}B_n^k+o(n^{k-1}) = (k+1)!\binom{n+k+1}{k}\phi+o(n^{k-1});$$

and this is (6.7.4), with $A = 0$, $h = (k+1)!\,\phi$, and $n+1$ for n. The proof is valid, and gives (6.7.5), for $k = 0$.

(iv) It is now easy to complete the proof of the theorem. In the first place, the condition is sufficient, by (ii). Secondly, if $\sum a_n$ is summable (C, k), b_n is any solution of (6.7.1), and $\mathrm{b}_n = b_n-h$, where h is the h of (6.7.4), then b_n is also a solution; and

$$\mathrm{B}_n^{k-1} = B_n^{k-1}-\binom{n+k}{k}h = \binom{n+k-1}{k-1}A+o(n^{k-1}),$$

by (6.7.4) or (6.7.5), so that $\sum \mathrm{b}_n$ is summable $(C, k-1)$ to sum A, the result holding for $k = 0$ since h is plainly independent of k. Hence the condition is also necessary.

Finally, since $\sum \mathrm{b}_n$ is summable $(C, k-1)$,

$$\mathrm{b}_n = \sum_n^\infty (\mathrm{b}_\nu-\mathrm{b}_{\nu+1}) = \sum_n^\infty \frac{a_\nu}{\nu+1} \quad (C, k-1),$$

by Theorem 48. For b_n, the h and ϕ of (iii) are 0, and (6.7.7) and (6.7.8) give

$$\mathrm{b}_0 = \mathrm{B}_0^k = (k+1)!\,\phi_0 = (k+1)!\sum \frac{A_\nu^k}{(\nu+1)(\nu+2)...(\nu+k+2)}.$$

Thus

$$\sum \frac{a_n}{n+1} = (k+1)!\sum \frac{A_n^k}{(n+1)(n+2)...(n+k+2)} \quad (C, k-1).$$

This is (6.6.3). Our proof here is independent of Theorems 71 and 76.

A corollary of Theorem 77 is

THEOREM 78. *A necessary and sufficient condition that* $\sum a_n$ *should be summable* (C, k) *is that there should be a system of numbers* $a_{n,s}$, *where* $s = 0, 1, \ldots, k+1$, *such that*

$$a_{n,0} = a_n, \qquad a_{n,s-1} = (n+1)(a_{n,s} - a_{n+1,s}) \quad (s > 0),$$

and that $$A_{k+1} = \sum a_{n,k+1}$$

is summable (C, -1). *In these circumstances*

$$a_{n,s} = \frac{a_{n,s-1}}{n+1} + \frac{a_{n+1,s-1}}{n+2} + \ldots \quad (C, k-s)$$

for $s = 1, 2, \ldots, k+1$; $A_s = \sum a_{n,s}$ *is summable* (C, $k-s$); *and the sums of all these series are the same.*

We have only to apply Theorem 77 $k+1$ times in succession. Theorems 77 and 78 may be used to obtain instructive proofs of the equivalence theorem (§ 5.8) and of other standard theorems in the subject.

6.8. Integrals. There are 'Tauberian' theorems for integrals like those of § 6.1. If

(6.8.1) $$\int a(x)\,dx = A \quad (C, k)\dagger$$

for some k, and $a(x) = O(x^{-1})$ for large x, then (6.8.1) is true for all $k > -1$: in particular, the integral is convergent. If (6.8.1) is true for some k, $a(x)$ is real, and $xa(x) > -H$, then the integral is convergent. The analogues of the preliminary Theorems 65 and 66 are: (i) if the integral is summable (C, $r+1$), then a necessary and sufficient condition for summability (C, r) is $B_r(x) = o(x^{r+1})$, where $B_r(x)$ is formed from $b(x) = xa(x)$ as $A_r(x)$ is formed from $a(x)$; and (ii) a necessary and sufficient condition for summability (C, $r+1$) is that $\int x^{-r-2}B_r(x)\,dx$ should be convergent.

There is an analogue of Theorem 71: if (6.8.1) is true, $f(x) \to 0$, $f^{(k)}(x)$, the kth derivative of $f(x)$, is absolutely continuous, and $\int x^k |f^{(k+1)}(x)|\,dx < \infty$, then

(6.8.2) $$\int a(x)f(x)\,dx = (-1)^{k+1}\int A_k(x)f^{(k+1)}(x)\,dx \quad (C, k),$$

the last integral being absolutely convergent. In particular this is true if $f(x) = (x+1)^{-s}$, where $s > 0$. On the other hand, there is no analogue of Theorem 76; the introduction of a convergence factor like $(x+1)^{-s}$ does not necessarily decrease the order of summability needed. Thus if $a(x) = e^{2x}\cos e^x$ then

$$A_1(x) = -Hx + \cos 1 - \cos e^x + \int_0^x \cos e^t\,dt \sim -Hx,$$

where $H = \cos 1 + \sin 1$, and so

(6.8.3) $$\int e^{2x}\cos e^x\,dx = -\cos 1 - \sin 1 \quad (C, 1).$$

But $$\int \frac{e^{2x}\cos e^x}{x+1}\,dx$$

is not convergent, and indeed not summable (C, k) for any $k < 1$.

† As in § 5.14, integrals written without limits are over $(0, \infty)$.

If

$$J(\delta) = \int e^{-\delta x} a(x)\, dx \tag{6.8.4}$$

is convergent for $\delta > 0$, and tends to A when $\delta \to 0$, then we write

$$J = \int a(x)\, dx = A \ \ (\mathrm{A}) \tag{6.8.5}$$

and say that J is *summable* (A) to A. It would be natural, after § 5.12, to expect summability (C, k) to involve summability A, but this is one of the points where the analogy between series and integrals breaks down. The summability of J does not involve even the convergence of $J(\delta)$. Thus the integral (6.8.3) is summable (C, 1), but $\int e^{(2-\delta)x} \cos e^x\, dx$ is not convergent if $\delta \leqslant 1$.

However, if J is summable (C, k), and $J(\delta)$ is convergent for every positive δ, then J is summable (A). For then

$$A^\delta(x) = \int_0^x e^{-\delta t} a(t)\, dt = O(1)$$

for every positive δ, and

$$A(x) = \int_0^x e^{\delta t} \frac{dA^\delta(t)}{dt}\, dt = e^{\delta x} A^\delta(x) - \delta \int_0^x e^{\delta t} A^\delta(t)\, dt = O(e^{\delta x})$$

for every such δ. It follows that $A_k(x) = O(e^{\delta x})$ for each k; and so, by $k+1$ partial integrations, that

$$J(\delta) = \int_0^\infty e^{-\delta x} a(x)\, dx = \delta^{k+1} \int_0^\infty e^{-\delta x} A_k(x)\, dx. \tag{6.8.6}$$

But $k!\, x^{-k} A_k(x) \to A$, and therefore

$$J(\delta) \sim A \frac{\delta^{k+1}}{k!} \int_0^\infty e^{-\delta x} x^k\, dx = A.$$

We can obtain a more satisfactory theorem as follows. The integral $\int x^k e^{-\delta x}\, dx$ is convergent for every k, so that, after (6.8.2), the summability (C, k) of J involves that of $J(\delta)$, and the truth of (6.8.6), the integral on the left being summable (C, k) and that on the right absolutely convergent. It then follows that $J(\delta)$, interpreted as a (C, k) integral, tends to A.

The integral $\int e^{aix + b\sqrt{x}}\, dx$, where $a > 0$, $b > 0$, is summable (A), but not (C, k) for any k.

6.9. The binomial series. In the rest of the chapter we study the summability of some particularly important special series. We begin with the series

$$\sum a_n = \sum \binom{n+\alpha}{\alpha} z^n,$$

where $\qquad \alpha = \beta + i\gamma, \qquad z = e^{i\theta}, \qquad |\theta| \leqslant \pi.$

It is familiar that the series is (1) absolutely convergent when $\beta < -1$,

(2) convergent, but not absolutely,† when $-1 \leqslant \beta < 0$ and $\theta \neq 0$, (3) divergent when $\beta \geqslant 0$, or when $-1 \leqslant \beta < 0$ and $\theta = 0$;† and that the sum of the series, when convergent, is

$$(1-z)^{-\alpha-1} = \exp\{-(\alpha+1)\log(1-z)\},$$

the logarithm having its principal value, for which $|\Im \log(1-z)| < \frac{1}{2}\pi$. We shall determine the conditions under which it is summable (C, k), for any $k > -1$.

We suppose first that

$$\beta \geqslant -1, \qquad k > -1, \qquad |\theta| \leqslant \pi, \qquad \theta \neq 0, \qquad z \neq 1.$$

Then

$$\sum a_n u^n = (1-zu)^{-\alpha-1}, \qquad \sum A_n^k u^n = (1-u)^{-k-1}(1-zu)^{-\alpha-1},$$

$$A_n^k = \frac{1}{2\pi i} \int\limits_C \frac{du}{(1-u)^{k+1}(1-zu)^{\alpha+1}u^{n+1}},$$

where $u = \rho e^{i\phi}$, C is the circle $\rho = \rho_0 < 1$, and the powers of $1-u$ and $1-zu$ have their principal values. Hence, by Cauchy's theorem,

$$A_n^k = \frac{1}{2\pi i}\left(\int\limits_{C_1} + \int\limits_{C_2}\right) \frac{du}{(1-u)^{k+1}(1-zu)^{\alpha+1}u^{n+1}} = J_1 + J_2,$$

where C_1 and C_2 are two contours surrounding the points $u = 1$ and $u = 1/z = \zeta$ and going to infinity in the directions $\phi = 0$ and $\phi = -\theta$ respectively. We may suppose C_1 and C_2 formed by circles round $u = 1$ and $u = \zeta$, and straight lines with arguments 0 and $-\theta$, these last described twice in opposite directions.

We write J_1 in the form $J_1 = J_1^{(1)} + J_1^{(2)}$, where

$$J_1^{(1)} = \frac{(1-z)^{-\alpha-1}}{2\pi i} \int\limits_{C_1} \frac{du}{(1-u)^{k+1}u^{n+1}} = \binom{n+k}{k}(1-z)^{-\alpha-1},\ddagger$$

$$J_1^{(2)} = \frac{1}{2\pi i} \int\limits_{C_1} \left\{\frac{1}{(1-uz)^{\alpha+1}} - \frac{1}{(1-z)^{\alpha+1}}\right\} \frac{du}{(1-u)^{k+1}u^{n+1}}.$$

We suppose that $n > |1-\zeta|^{-1}$ and take the radius of the circular part of C_1 to be n^{-1}, so that $u^{-n-1} = O(1)$ on the circle. Also

$$(1-uz)^{-\alpha-1} - (1-z)^{-\alpha-1} = (\alpha+1)z \int\limits_1^u (1-wz)^{-\alpha-2}\,dw$$

is $O(|u-1|)$ on the whole of C_1, and $O(n^{-1})$ on the circle. Thus the

† Except in the trivial case $\beta = -1$, $\gamma = 0$, $\alpha+1 = 0$, when the series reduces to its first term 1.

‡ Evaluating the integral by deforming C_1 back into C.

contribution of the circle to $J_1^{(2)}$ is $O(n^{-1}.n^{-1}.n^{k+1}) = O(n^{k-1})$, and that of the rest of C_1 is

$$O\left\{\int\limits_{1+n^{-1}}^{\infty} \frac{x-1}{(x-1)^{k+1}} x^{-n-1}\,dx\right\} = O\left\{\int\limits_{n^{-1}}^{\infty} \frac{t}{t^{k+1}} \frac{dt}{(1+t)^{n+1}}\right\}$$

$$= O\left\{n^{k+1}\int\limits_{0}^{\infty} \frac{t\,dt}{(1+t)^{n+1}}\right\} = O\left\{\frac{n^{k+1}}{n(n-1)}\right\} = O(n^{k-1}).$$

Hence
$$J_1 = \binom{n+k}{k}(1-z)^{-\alpha-1}+O(n^{k-1}).$$

Similarly we write $J_2 = J_2^{(1)}+J_2^{(2)}$, where

$$J_2^{(1)} = \frac{(1-\zeta)^{-k-1}}{2\pi i}\int\limits_{C_2} \frac{du}{(1-zu)^{\alpha+1}u^{n+1}}$$

$$= \binom{n+\alpha}{\alpha}\frac{z^n}{(1-\zeta)^{k+1}} = \binom{n+\alpha}{\alpha}\frac{e^{ni\theta}}{(1-e^{-i\theta})^{k+1}},$$

$$J_2^{(2)} = \frac{1}{2\pi i}\int\limits_{C_2}\left\{\frac{1}{(1-u)^{k+1}}-\frac{1}{(1-\zeta)^{k+1}}\right\}\frac{du}{(1-zu)^{\alpha+1}u^{n+1}};$$

and we can prove that $J_2^{(2)} = O(n^{\beta-1})$ by an argument like that which we used for $J_1^{(2)}$.

Collecting our results, we find that

$$A_n^k = \binom{n+k}{k}(1-e^{i\theta})^{-\alpha-1}+O(n^{k-1})+\binom{n+\alpha}{\alpha}e^{ni\theta}(1-e^{-i\theta})^{-k-1}+O(n^{\beta-1}).$$

The first term here is the dominating term when $k > \beta$, the third when $k < \beta$; if $k = \beta$ then these terms are of the same order of magnitude. We thus obtain

THEOREM 79. *If* $\alpha = \beta+i\gamma$, $\beta \geqslant -1$, $k > -1$, $|\theta| \leqslant \pi$, *and* $\theta \neq 0$, *then the series* $\sum \binom{n+\alpha}{\alpha} e^{ni\theta}$ *is summable* (C, k) *when* $k > \beta$, *to sum* $(1-e^{i\theta})^{-\alpha-1}$. *It oscillates finitely* (C, k) *when* $k = \beta$, *and infinitely when* $k < \beta$.

It is plain that the argument will prove uniform summability in any closed interval of θ which does not include $\theta = 0$.

When $\theta = 0$, $z = 1$, $a_n = \binom{n+\alpha}{\alpha}$,

$$\sum A_n^k u^n = (1-u)^{-\alpha-k-2}, \qquad A_n^k = \binom{n+\alpha+k+1}{\alpha+k+1}$$

and
$$\binom{n+k}{k}^{-1} A_n^k \sim \frac{\Gamma(k+1)}{\Gamma(\alpha+k+2)}n^{\alpha+1}.$$

Thus we obtain

THEOREM 80. *The series* $\sum \binom{n+\alpha}{\alpha}$ *is never summable* (C, k), *for any* k, *unless* $\beta < -1$, *in which case it converges absolutely to* 0, *or* $\beta = -1$, $\gamma = 0$, *in which case it reduces to its first term* 1.

6.10. The series $\sum n^{\alpha} e^{ni\theta}$. From Theorems 79 and 80 we can deduce corresponding results for the series $\sum n^{\alpha} e^{ni\theta}$.† We suppose for the present that $\beta > -1$, leaving the case $\beta = -1$ to the next section.

If $\gamma = 0$ and β is integral (so that α is integral) then

$$n^{\alpha} = p_0\binom{n+\alpha}{\alpha}+p_1\binom{n+\alpha-1}{\alpha-1}+\dots+p_{\alpha}, \tag{6.10.1}$$

where $p_0, p_1,\dots$ are independent of n. If α is not an integer then

$$\binom{n+\alpha-\nu}{\alpha-\nu} = c_{\nu,\nu} n^{\alpha-\nu}+c_{\nu,\nu+1} n^{\alpha-\nu-1}+\dots+c_{\nu,h} n^{\alpha-h}+O(n^{\beta-h-1}),$$

where h is arbitrary, $\nu = 0, 1,\dots, h$, and $c_{\nu,\nu} \neq 0$; and, combining these equations, we can express n^{α} in the form

$$n^{\alpha} = p_0\binom{n+\alpha}{\alpha}+p_1\binom{n+\alpha-1}{\alpha-1}+\dots+p_h\binom{n+\alpha-h}{\alpha-h}+O(n^{\beta-h-1}). \tag{6.10.2}$$

Combining Theorem 79 with (6.10.1) or (6.10.2), we obtain

THEOREM 81. *If* $\alpha = \beta+i\gamma$, $\beta > -1$, $k > -1$, $|\theta| \leqslant \pi$ *and* $\theta \neq 0$, *then the series* $\sum n^{\alpha}e^{ni\theta}$ *is summable* (C, k) *when* $k > \beta$, *oscillates finitely* (C, k) *when* $k = \beta$, *and oscillates infinitely* (C, k) *when* $k < \beta$.

6.11. The case $\beta = -1$. The case in which $\beta = -1$ is in some ways particularly interesting. The series

$$\sum \binom{n-1+i\gamma}{-1+i\gamma} e^{ni\theta}, \qquad \sum n^{-1+i\gamma} e^{ni\theta}, \tag{6.11.1}$$

where $\gamma \neq 0$, are convergent unless $\theta \equiv 0 \pmod{2\pi}$. In that case they oscillate finitely, since

$$\binom{n-1+i\gamma}{-1+i\gamma} = \frac{\Gamma(n+i\gamma)}{\Gamma(i\gamma)\Gamma(n+1)} = \frac{n^{-1+i\gamma}}{\Gamma(i\gamma)}+O\left(\frac{1}{n^2}\right),$$

$$\sum_1^{n-1} m^{-1+i\gamma}-\frac{n^{i\gamma}-1}{i\gamma}$$

$$= \sum_1^{n-1} m^{-1+i\gamma}-\int_1^n t^{-1+i\gamma}\,dt = \sum_1^{n-1}\int_m^{m+1} (m^{-1+i\gamma}-t^{-1+i\gamma})\,dt,$$

† The series starting from $n = 1$ when $\beta \leqslant 0$.

and the general term of the last sum is $O(m^{-2})$.† Since the series (6.11.1) are not convergent, and their general terms are $O(n^{-1})$, it follows from Theorem 63 that they are not summable (C, k) for any k.

It is interesting to investigate certain other properties of these series, and in particular to prove their partial sums bounded uniformly in θ. More generally, we prove

THEOREM 82. *If*

(6.11.2) $$A_n = a_0+a_1+...+a_n = O(1), \qquad \Delta a_n = O(n^{-2}),$$

then

(6.11.3) $$|s_n(z)| = \left|\sum_0^n a_m z^m\right| < H$$

for $|z| \leqslant 1$, *H being independent of z and n. In particular this is true when* a_n *is* $\binom{n-1+i\gamma}{-1+i\gamma}$ *or* $n^{-1+i\gamma}$ *(with* $a_0 = 0$ *in the second case).*

We note in passing that the hypotheses (6.11.2) imply $a_n = O(n^{-1})$. In what follows we shall be dealing with functions of n and z, and O's will be uniform for $|z| \leqslant 1$. It is sufficient, by the principle of the maximum modulus, to prove (6.11.3) when $z = e^{i\theta}$ and $0 < |\theta| \leqslant \pi$. We suppose $\theta > 0$, and write $p = [\pi/\theta]$, so that $p \geqslant 1$.

If $p \geqslant n$ then

$$s_n(z) = \sum_0^n a_m e^{mi\theta} = \sum_0^{n-1} A_m \Delta e^{mi\theta}+A_n e^{ni\theta}$$
$$= (1-e^{i\theta})\sum_0^{n-1} O(1)+O(1) = O(n\theta)+O(1) = O(1).$$

If $p < n$ then

$$s_n(z) = \sum_0^p a_m e^{mi\theta}+\sum_{p+1}^n a_m e^{mi\theta} = S_1+S_2,$$

and the argument just used shows that $S_1 = O(1)$. Also

$$(1-e^{i\theta})S_2 = \sum_{p+1}^n a_m \Delta e^{mi\theta} = a_{p+1} e^{(p+1)i\theta}-a_n e^{(n+1)i\theta}-\sum_{p+2}^n e^{mi\theta}\Delta a_{m-1}$$
$$= O(p^{-1})+\sum_{p+1}^n O(m^{-2}) = O(p^{-1}) = O(\theta),$$

since $a_n = O(n^{-1})$, and so $S_2 = O(1)$.

It follows from Theorem 82 that, for example, the series

(6.11.4) $$\sum \frac{n^{-1+i\gamma}}{\log n} z^n \quad (n = 2, 3,..., \gamma \neq 0)$$

is uniformly, though not absolutely, convergent on the unit circle.

† Alternatively, we may use the identity $\sum_0^n \binom{m-1+i\gamma}{-1+i\gamma} = \binom{n+i\gamma}{i\gamma}$.

We leave it to the reader to prove

THEOREM 83. *If* $-1<\beta<0$, $a_n=o(1)$, $\Delta a_n=O(n^{\beta-1})$, *then*

$$|s_n(z)|<H|1-z|^{-\beta-1} \tag{6.11.5}$$

for $|z|\leqslant 1$. *If* $a_n=O(1)$, $\Delta^2 a_n=O(n^{-2})$, *then* (6.11.5) *is true for* $\beta=0$.

6.12. The series $\sum n^{-b}e^{Ain^a}$. The series

$$\sum a_n=\sum n^{-b}e^{Ain^a}, \tag{6.12.1}$$

where $\qquad A>0, \qquad 0<a<1, \qquad b=\beta+i\gamma,$

is particularly interesting and may be used to illustrate many points in the theory of summable series. It is absolutely convergent if $\beta>1$, and we shall suppose throughout that $\beta\leqslant 1$.

The order of a_n is decreased, by a factor n^{a-1}, by differentiation, so that the series is adapted for study by means of the Euler-Maclaurin sum formula; but the discussion of its summability on these lines is rather tiresome in detail, and we shall use a different method depending on a direct use of Cauchy's theorem.

THEOREM 84. *The series* (6.12.1) *is summable* (C, k), *where* $k>-1$, *if and only if*

$$(k+1)a+\beta>1. \tag{6.12.2}$$

We write

$$u(z)=z^{-b}e^{Aiz^a}, \qquad u_0=0, \qquad u_n=u(n) \quad (n>0),$$

z^{-b} and z^a having their principal values in the half plane $\Re(z)>0$, and

$$S=\Gamma(k+1)U_n^k=\sum_{m=1}^{n}\frac{\Gamma(n-m+k+1)}{\Gamma(n-m+1)}u_m. \tag{6.12.3}$$

We have to show that $n^{-k}S$ tends to a limit if and only if k satisfies (6.12.2).

We denote by C the rectangle $(\frac{1}{2}-iY, n-iY, n+iY, \frac{1}{2}+iY)$, shown in Fig. 1, by C_1 and C_2 the two half-rectangles formed by the lines L_1 to L_4 and L_5 to L_8 respectively. If

$$f(z)=\frac{\Gamma(n-z+k+1)}{\Gamma(n-z+1)}u(z),$$

then Cauchy's theorem gives

$$S'=S-\tfrac{1}{2}f(n)=\frac{1}{2\pi i}\int_C \pi\cot\pi z f(z)\,dz, \tag{6.12.4}$$

where the integral along $(n-iY,\ n+iY)$ is a principal value.† Also

$$(6.12.5)\qquad \frac{1}{2\pi i}\int\limits_{C_1} \pi i f(z)\,dz = 0, \qquad \frac{1}{2\pi i}\int\limits_{C_2} (-\pi i) f(z)\,dz = 0.$$

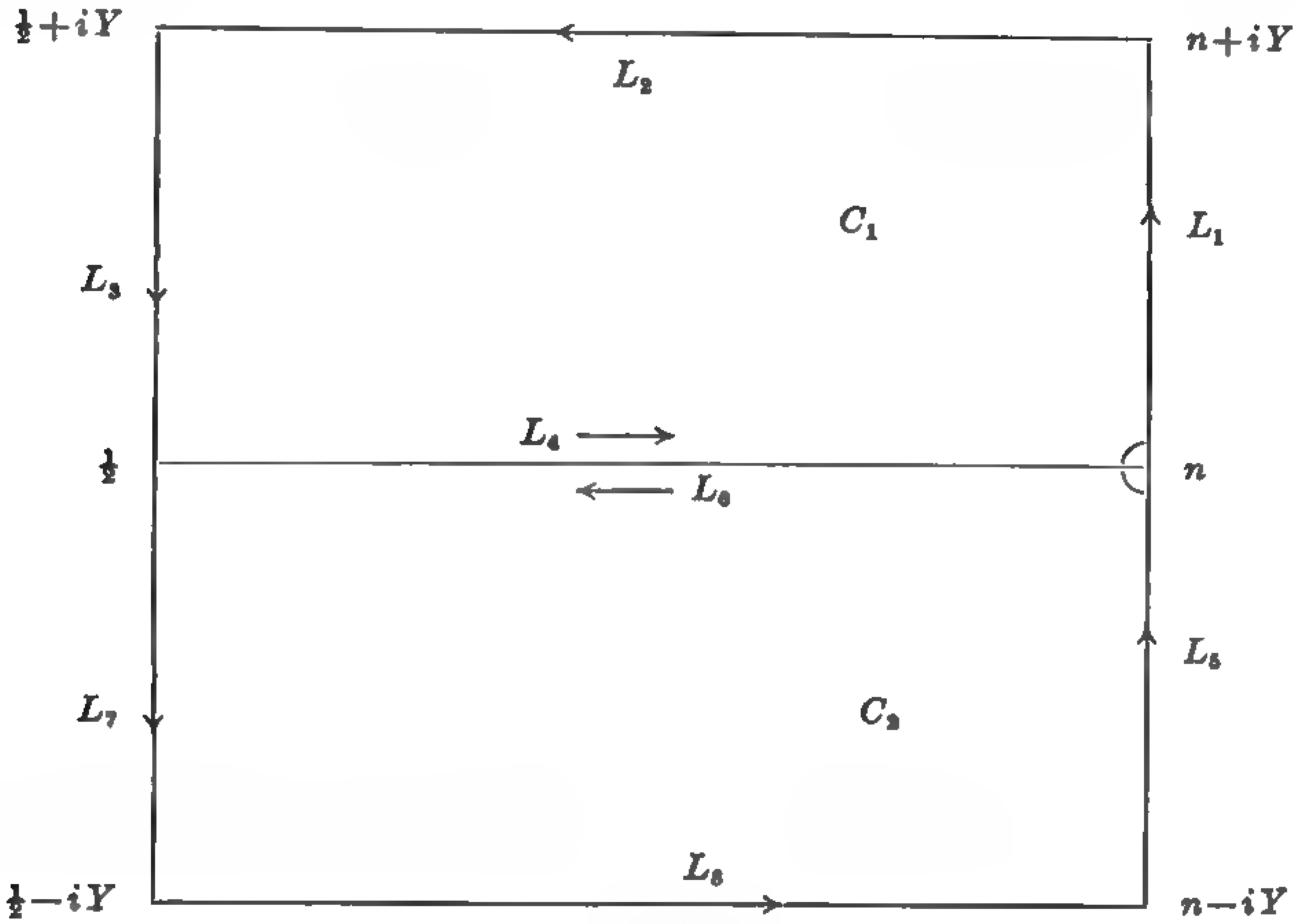

Fig. 1.

Hence, combining (6.12.4) and (6.12.5), we obtain

$$(6.12.6)\qquad S' = \frac{1}{2\pi i}\int\limits_{L_2} \psi(z)f(z)\,dz + \frac{1}{2\pi i}\int\limits_{L_8} \psi(z)f(z)\,dz +$$

$$+\int\limits_{\frac{1}{2}}^{n} f(z)\,dz + \frac{1}{2\pi i}\int\limits_{L_3+L_7} \psi(z)f(z)\,dz + \frac{1}{2\pi i}\int\limits_{L_5+L_1} \psi(z)f(z)\,dz,$$

where
$$\psi(z) = \pi(\cot \pi z \pm i)$$

according as $y = \Im(z)$ is positive or negative. The integral along L_5+L_1 is a principal value.

Now

$$\psi(z) = O(e^{-2\pi|y|}), \qquad \frac{\Gamma(n-z+k+1)}{\Gamma(n-z+1)} = O(|y|^k), \qquad e^{Aiz^a} = O(e^{A|y|^a})$$

† We apply Cauchy's theorem to C modified by a semicircular indentation round $z = n$, and then make the radius of the indentation tend to zero.

for fixed n, $\frac{1}{2} \leqslant x \leqslant n$, and large $|y|$. Hence the integrals along L_2 and L_3 tend to 0 when $Y \to \infty$, and (6.12.6) gives

$$(6.12.7) \qquad S' = \int_{\frac{1}{2}}^{n} f(z)\,dz - \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi(\tfrac{1}{2}+iy) f(\tfrac{1}{2}+iy)\,dy + \\ + \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi(n+iy) f(n+iy)\,dy = J - J_1 + J_2,$$

say, the last integral being a principal value.

The integral

$$\int_{-\infty}^{\infty} \left\{ n^{-k} \frac{\Gamma(n+\frac{1}{2}+k-iy)}{\Gamma(n+\frac{1}{2}-iy)} \right\} u(\tfrac{1}{2}+iy)\psi(\tfrac{1}{2}+iy)\,dy$$

is majorized by an integral with integrand independent of n, and the function in curly brackets tends to 1 when $n \to \infty$, for every y. Hence

$$(6.12.8) \qquad n^{-k} J_1 \to \frac{1}{2\pi} \int_{-\infty}^{\infty} u(\tfrac{1}{2}+iy)\psi(\tfrac{1}{2}+iy)\,dy \\ = i \int_{-\infty}^{\infty} (\tfrac{1}{2}+iy)^{-b} e^{Ai(\frac{1}{2}+iy)^a} \frac{\operatorname{sgn} y}{e^{2\pi|y|}+1}\,dy = iI,$$

say. As regards J_2, we have

$$\psi(n+iy) = \frac{2\pi i}{1-e^{2\pi y}} \quad (y > 0), \qquad -\frac{2\pi i}{1-e^{-2\pi y}} \quad (y < 0);$$

and so

$$(6.12.9) \\ J_2 = -i \int_{-\infty}^{\infty} \frac{\Gamma(k+1-iy)}{\Gamma(1-iy)} u(n+iy) \frac{\operatorname{sgn} y}{e^{2\pi|y|}-1}\,dy \\ = -i \int_{0}^{\infty} \left\{ \frac{\Gamma(k+1-iy)}{\Gamma(1-iy)} u(n+iy) - \frac{\Gamma(k+1+iy)}{\Gamma(1+iy)} u(n-iy) \right\} \frac{dy}{e^{2\pi y}-1},$$

the last integral being absolutely convergent. It follows that

$$J_2 = O(n^{-\beta}),\dagger$$

† We divide the range of integration into $(0, \delta)$ and (δ, ∞), where $0 < \delta < 1$. It is clear, on grounds of 'dominated convergence', that the part of the integral over (δ, ∞) is $O(n^{-\beta})$. In $(0, \delta)$ we may expand

$$n^b u(n+iy), \qquad \frac{\Gamma(k+1-iy)}{\Gamma(1-iy)}, \qquad \frac{y}{e^{2\pi y}-1}$$

as uniformly convergent power series $P(y/n)$, $Q(y)$, $R(y)$; and it then becomes plain that

$$n^{-b} \int_{0}^{\delta} \frac{1}{y} \left\{ P\left(\frac{y}{n}\right) Q(y) - P\left(-\frac{y}{n}\right) Q(-y) \right\} R(y)\,dy = O(n^{-\beta}).$$

and from (6.12.7) and (6.12.8) that

(6.12.10) $$S' = -in^k I + o(n^k) + J + O(n^{-\beta}).$$

Now

(6.12.11) $$J = \int_{\frac{1}{2}}^{n} \frac{\Gamma(n-z+k+1)}{\Gamma(n-z+1)} u(z)\, dz = \int_{M_1} - \int_{M_2} = I_1 - I_2,$$

where M_1 and M_2 are the lines through $\frac{1}{2}$ and n parallel to the positive direction of the imaginary axis. On M_1, $z = re^{i\theta}$, where $0 < \theta < \frac{1}{2}\pi$, $\theta \to \frac{1}{2}\pi$; and $|e^{Aiz^a}| = e^{-Ar^a \sin a\theta}$. It follows† that

(6.12.12) $$I_1 \sim in^k \int (\tfrac{1}{2}+iy)^{-b} e^{Ai(\frac{1}{2}+iy)^a}\, dy = in^k I^*.$$

We now take a (small) fixed δ, and write

(6.12.13)
$$I_2 = i \int_0^{\infty} \frac{\Gamma(k+1-iy)}{\Gamma(1-iy)} u(n+iy)\, dy = i \int_0^{\delta n} + i \int_{\delta n}^{\infty} = I_3 + I_4.$$

In I_4, $y > \delta n$, and $z = n+iy = re^{i\theta}$, where $0 < \omega < \theta < \frac{1}{2}\pi$ and ω depends only on δ. Hence

$$|e^{Aiz^a}| = e^{-Ar^a \sin a\theta} < e^{-Br^a},$$

$$|z^{-b}| = r^{-\beta} e^{\gamma\theta} < Cr^{-\beta}, \qquad dy = \operatorname{cosec}\theta\, dr < D\, dr,$$

and

(6.12.14) $$I_4 = O\left(\int_{\delta n}^{\infty} r^{k-\beta} e^{-Br^a}\, dr\right) = O(e^{-En^a}),$$

B, C, D, and E being positive functions of δ.

Finally, if $0 < y < \delta n$, we have

$$e^{Ai(n+iy)^a} = e^{Ain^a - Aan^{a-1}y + \dots} = O(e^{-Fn^{a-1}y}),$$

where F is a positive function of δ. Hence

(6.12.15) $$I_3 = O\left\{n^{-\beta}\left(\int_0^1 e^{-Fn^{a-1}y}\, dy + \int_1^{\delta n} y^k e^{-Fn^{a-1}y}\, dy\right)\right\}$$

$$= O(n^{-\beta}) + O\left(n^{-\beta} \int_0^{\infty} y^k e^{-Fn^{a-1}y}\, dy\right) = O\{n^{-\beta+(k+1)(1-a)}\},$$

since $(k+1)(1-a) > 0$.

Finally, S and S' differ by $\frac{1}{2}f(n) = O(n^{-\beta})$. Hence, collecting our results from (6.12.10)–(6.12.15), and remembering again that

$$(k+1)(1-a) > 0,$$

we find that

$$S = in^k(I^* - I) + o(n^k) + O\{n^{-\beta+(k+1)(1-a)}\}.$$

† Again by a simple argument based on majorization.

If k satisfies (6.12.2) then $-\beta+(k+1)(1-a) < k$ and $n^{-k}S \to i(I^*-I)$, so that the series (6.12.1) is summable (C, k), to sum $i(I^*-I)$.

To prove the negative assertion of the theorem, we must estimate I_2 more precisely. If we replace $(n+iy)^{-b}$ by n^{-b}, and $Ai(n+iy)^a$ by $Ain^a-Aan^{a-1}y$, we obtain

$$in^{-b}e^{Ain^a}\int \frac{\Gamma(k+1-iy)}{\Gamma(1-iy)}e^{-Aan^{a-1}y}\,dy.$$

When n is large, n^{a-1} is small, so that the integral here is dominated by the part in which y is large; and we are thus led to replace the quotient of gamma-functions by $(-iy)^k$. This leads us to the conclusion that

$$-I_2 \sim -in^{-b}e^{Ain^a}\int (-iy)^k e^{-Aan^{a-1}y}\,dy$$
$$= e^{-\frac{1}{2}(k+1)\pi i}\Gamma(k+1)(Aa)^{-k-1}n^{-b+(k+1)(1-a)}e^{Ain^a}.$$

There is no particular difficulty in making this conclusion rigorous, and we suppress the details of the proof. It follows that, when $(k+1)a+\beta \leqslant 1$, S involves an oscillating term whose order is at least n^k, and that the series is not summable.

NOTES ON CHAPTER VI

§ 6.1. Theorem 63 was proved by Hardy, *PLMS* (2), 8 (1910), 301–20, except for the clause concerning summability by means of negative order, which was added by Hardy and Littlewood (l.c. under § 5.7); and Theorem 64 by Landau, *PMF*, 21 (1910), 97–177 (103–13). The method of proof here, based on Theorems 65 and 66, is Hardy's. A good many other proofs have been given, particularly for the special case $k = 1$, which is important in the theory of Fourier series. See, for example, Bromwich, 423–6; de la Vallée-Poussin, *Cours d'analyse infinitésimale* (éd. 6, Louvain, 1926, ii, 109); Kloosterman, *JLMS*, 15 (1940), 91–6; Mordell, *JLMS*, 3 (1928), 86–9, 119–21, 170–2.

Theorem 67 was found by Hardy, *PLMS* (2), 12 (1913), 174–80, in the more general form in which $\sum a_n$ is given summable by Riesz's typical means of *some* order. A gap in Hardy's proof was filled by Ananda Rau, *PLMS* (2), 17 (1918), 334–6. The form of the proof here for $k = 1$ is due to Bosanquet.

Hardy states erroneously that $a_n > -H(\lambda_n-\lambda_{n-1})/\lambda_n$ is a sufficient condition: the mistake was corrected by Ananda Rau, *PLMS* (2), 30 (1930), 367–72. On the other hand, the two conditions

$$(1)\quad a_n > -H(\lambda_n-\lambda_{n-1})/\lambda_n, \qquad (2)\quad \varliminf a_n \geqslant 0$$

are sufficient: in this case $A(x)$ is slowly decreasing in the sense of § 6.2. This theorem is included in one due to Szász, *Münchener Sitzungsberichte* (1929), 325–40: see the note on § 7.7. If $\lambda_{n+1}/\lambda_n \to 1$, then (1) implies (2) and is sufficient in itself.

§ 6.2. The definitions of slowly oscillating and slowly decreasing functions and sequences are due to R. Schmidt, *MZ*, 22 (1924), 89–152 (127–42). We shall use two forms of the definitions, the first appropriate to the interval $(0, \infty)$, the second to $(-\infty, \infty)$: see § 12.2. It is the first form which is relevant here.

§ 6.3. Hardy and Littlewood, *MM*, 43 (1914), 134–47. Actually the convergence of $\sum n^{p-1}|a_n|^p$ is a sufficient condition for the corresponding theorem concerning summability (A).

§ 6.4. Theorem 70, for integral parameters, is proved (though not quite explicitly) by Hardy and Littlewood, *PLMS* (2), 11 (1913), 411–78 (437). The theorem is the case $\beta = 0$ of their Theorem 19, with 'bounded $(C, r-k)$, summable (C, r)' in the hypotheses. They state their result only for $\beta > 0$, but the proof is valid for $\beta = 0$.

There is a considerable literature concerning the general form of the theorem with unrestricted parameters, and extensions of it important in the theory of Dirichlet's series. See, for example, Ananda Rau, *PLMS* (2), 34 (1932), 414–40; Andersen, *Studier*, 55 et seq.; Bosanquet, *JLMS*, 18 (1943), 239–48; M. Riesz, *MTE*, 29 (1911), 283–301, and *AUH*, 1 (1923), 104–13; Zygmund, *MZ*, 25 (1926), 291–6. Bosanquet gives further references.

§§ 6.5–6. Theorem 71 was proved independently by Bohr [*CR*, 148 (1909), 75–80; *Bidrag*, 61–9] and by Hardy [*PLMS* (2), 6 (1908), 255–64; and 8 (1910), 277–94 (278–81), where a mistake in the earlier paper is corrected]. A number of special cases had been proved earlier by various writers, e.g. by Bromwich, *MA*, 65 (1908), 350–69, and by Hardy [*PLMS* (2), 4 (1906), 247–65 and *MA*, 64 (1907), 77–94].

The theorem was extended to general k by Andersen, *Studier*, 44–55. Simplified proofs of the generalized theorem, and further extensions, have been given by Andersen, *PLMS* (2), 27 (1928), 39–71, and Bosanquet, *JLMS*, 17 (1942), 166–73.

The necessity of the conditions (in the sense explained on pp. 130–1) was proved for integral k by Fekete, *MTE*, 35 (1917), 309–24, and for general k by Bosanquet, l.c. *supra*.

There are a number of theorems which include both of Theorems 71 and 76, especially for integral parameters. Thus Bosanquet, *PLMS* (2), 50 (1948), 295–304, has proved that *if k and l are integers, $-1 \leqslant l \leqslant k$, and p is any real number, then, in order that $\sum a_n f_n$ should be summable (C, l) whenever $A_n^k = O(n^{k+p})$, it is necessary and sufficient that*

$$f_n = o(n^{l-p-k}), \qquad \sum n^{p+k}|\Delta^{k+1} f_n| < \infty.$$

If, for example, $p = 0$, we obtain necessary and sufficient conditions that $\sum a_n f_n$ should be summable (C, l) whenever $\sum a_n$ is summable or bounded (C, k). This case of the theorem was stated without proof by Schur, *JM*, 151 (1921), 79–111 (106), and proved by Bosanquet, *JLMS*, 20 (1945), 39–48. It reduces to Theorem 71 for $l = k$; and there is a variant in which $\sum a_n$ is summable (C, k) and $f_n = O(n^{l-k})$.

The special case $l = 0$, $p = 0$ is considerably older. The sufficiency of the conditions in this case was proved by Bromwich, l.c. *supra*, for integral k, and by Chapman, l.c., under § 5.5, generally; and the necessity by Kojima, *TMJ*, 12 (1917), 291–326. See Moore, *Convergence factors*, 45–6.

More recently Bosanquet, *PLMS* (not yet published), has extended his theorem, with the slightly narrower conditions $0 \leqslant l \leqslant k$, $p \geqslant 0$, to non-integral k and l.

Theorem 76 was stated (at any rate for integral k) by M. Riesz, *CR*, 148 (1909), 1658–60; and proved, for general k, integral s, by Chapman, l.c., under § 5.5, 388–9. There is a proof for general k and s by Zygmund, *BAP* (1927), 309–31; and another by Ananda Rau, left incomplete at one point in his paper

referred to under § 6.4, has been completed by Minakshisundaram, *JIMS* (2), 2 (1936), 147–55.

There are analogues of these theorems for 'absolute summability'. The series $\sum a_n$ is said to be absolutely summable (C, k), or summable $|C, k|$, if

$$\sum |C_n^k(A) - C_{n-1}^k(A)| < \infty.$$

Thus summability $|C, 0|$ is absolute convergence: the definition is Fekete's. We shall not be concerned with absolute summability here, but theorems corresponding to Theorem 71, and others of these sections, have been proved by Bosanquet, Fekete, and Kogbetliantz. References will be found in Kogbetliantz and in Bosanquet's papers quoted here.

We add a remark about the argument in the text. We deduce Theorem 76 (for integral k and s) from Theorems 47, 71, and 65. Alternatively, we may deduce it from Theorems 47 and 66. It is trivial if $k = 0$. If $k > 0$ then, in order that $\sum n^{-1}a_n$ should be summable (C, $k-1$), it is necessary and sufficient that $\sum n^{-k}A_n^{k-1}$ should be convergent, and this is easily proved by partial summation. The proof is valid for non-integral k.

We can also vary the proof so as to avoid an appeal to Theorem 47.

§ 6.7. Theorems 77 and 78 were proved by Hardy and Littlewood, *MZ*, 19 (1924), 67–96: they are closely connected with others proved independently by Knopp, ibid. 97–113. Later, Andersen, *PLMS* (2), 27 (1928), 39–71, and Hardy and Littlewood, ibid. 327–48, transformed and generalized them in various ways. See Kogbetliantz, 33.

§ 6.8. It is, as usual, difficult to give precise references for the integral theorems. For the equivalence theorem, see Landau, *Leipziger Sitzungsberichte*, 65 (1913), 131–8; for the analogue of Theorem 71, Hardy, *MM*, 40 (1910), 108–12; for the points discussed at the end of the section, M. E. Grimshaw, *JLMS*, 9 (1934), 94–102.

§§ 6.9–10. The substance of the results here is due to Chapman and Knopp, l.c., under § 5.5.

§ 6.11. The bounded convergence of the series (6.11.1), and the uniform convergence of (6.11.4), were proved, less directly, by Hardy, *QJM*, 44 (1913), 147–60. See also Landau, *Ergebnisse*, 68–9.

The most interesting case of Theorem 83, in which $\sum a_n z^n = (1-z)^{-\beta-1}$, is equivalent to a theorem of M. Riesz, *AUH*, 1 (1923), 114–26. It is stated more explicitly by Fejér, *MZ*, 24 (1925), 267–84 (269). Szegö, *MZ*, 25 (1926), 172–87, gives a different proof, based on Kaluza's Theorem 22, and a generalization to the case $\beta > 0$.

The proof of Theorem 83 is a little more complex when $\beta = 0$ than when $\beta < 0$. It is worked out in detail for the case $a_n = n^{i\gamma}$ by Hardy and Rogosinski, *OQJ*, 16 (1945), 49–58.

§ 6.12. The main result is due to Hardy, *PLMS* (2), 9 (1911), 126–44; but the discussion there is not quite satisfactory for our present purpose, since it is based on the restricted form of Riesz's means of § 5.16 in which ω assumes integral values only. It is not difficult to modify the argument so as to take account of non-integral ω, and prove that the series is summable (R, n, k) when $(k+1)a+\beta > 1$; but then we need the troublesome Theorem 58 (proved in § 5.16 only for integral k) in order to infer summability (C, k).

VII

TAUBERIAN THEOREMS FOR POWER SERIES

7.1. Abelian and Tauberian theorems. We shall be concerned throughout this chapter with a set of theorems of the kind usually called 'Tauberian'. We used this word in § 6.1, and gave a short explanation of the nature of a Tauberian theorem. The theorems which we prove here are more difficult, and our exposition of them more systematic, so that it will be best to begin by a more precise definition of the meaning of the word and of the word 'Abelian' with which it is contrasted. It is convenient to use notations differing in some points from those which we have used hitherto.

We denote the series and integral

$$\sum a_n, \qquad \int a(t)\,dt \tag{7.1.1}$$

by S and J, and their values, when they are convergent, by s and j (so that, for example, $S = s$ means that $\sum a_n$ converges to s). We write

$$s_n = a_0+a_1+\ldots+a_n, \qquad j(t) = \int_0^t a(u)\,du,$$

and

$$S(y) = \sum a_n e^{-ny}, \qquad J(y) = \int a(t)e^{-yt}\,dt,$$

when $y > 0$ and the series and integral are convergent. By $S = s\,(\mathrm{A})$ or $J = j\,(\mathrm{A})$ we mean that $S(y) \to s$ or $J(y) \to j$ when $y \to 0$, and by $S = s\,(\mathrm{C})$ or $J = j\,(\mathrm{C})$ we mean that the series or integral (7.1.1) is summable (C, 1) to s or j: we shall not have occasion to consider Cesàro summability of any other order. We denote the hypotheses

$$S = s,\quad J = j,\quad S = s\ (\mathrm{A}),\quad J = j\ (\mathrm{A}),\quad S = s\ (\mathrm{C}),\quad J = j\ (\mathrm{C})$$

by

$$K,\ K',\ K_A,\ K'_A,\ K_C,\ K'_C$$

respectively.

An 'Abelian' theorem is, roughly, one which asserts that, if a sequence or function behaves regularly, then some average of the sequence or function behaves regularly. Thus 'if $s_n \to s$ then

$$\sigma_n = \frac{s_0+s_1+\ldots+s_n}{n+1} \to s\text{'}$$

or 'K implies K_C' and its integral analogue 'K' implies K'_C' are Abelian.

Abel's theorem on the continuity of power series is an Abelian theorem (and it is from this that the name is derived). For

$$a_0+a_1x+a_2x^2+\ldots = \frac{s_0+s_1x+s_2x^2+\ldots}{1+x+x^2+\ldots}$$

when $0 < x < 1$ and the series are convergent; the right-hand side is a certain average of the s_n; and Abel's theorem asserts that this average tends to s, when $x \to 1$, if s_n itself tends to s. Generally, any theorem asserting the regularity (§ 3.2) of a method of summation is an Abelian theorem.

The direct converse of an Abelian theorem is usually false. It is obvious, for example, that, if the regularity theorem for any method of summation is reversible, then the method is trivial in the sense that it will sum convergent series only. There are, however, many important theorems which may be called *corrected forms of the false converses of Abelian theorems*. Thus we saw in § 6.1 that the false theorem '$\sigma_n \to s$ implies $s_n \to s$', or 'K_C implies K', becomes true if we subject s_n to an appropriate additional condition, such as $a_n = O(n^{-1})$. Such theorems are called 'Tauberian', after A. Tauber, who first proved one of the simplest of them; and the supplementary condition is called a 'Tauberian condition'.

The most important Tauberian conditions with which we shall be concerned here are

$$(o)\quad a_n = o(n^{-1}),\qquad (O)\quad a_n = O(n^{-1}),\qquad (O_L)\quad a_n > -Hn^{-1},$$
$$(O_R)\quad a_n < Hn^{-1},$$

and their integral analogues

$$(o')\quad a(t) = o(t^{-1}),\qquad (O')\quad a(t) = O(t^{-1}),\qquad (O'_L)\quad a(t) > -Ht^{-1},$$
$$(O'_R)\quad a(t) < Ht^{-1}.$$

Here H is a positive constant, and the conditions on $a(t)$ are supposed to be satisfied for large t. The behaviour of $a(t)$ for *small* t will be irrelevant; we shall usually suppose only that it is integrable down to 0.

We shall also use two generalizations of (o) and (o'), viz.

$$(\omega)\qquad a_1+2a_2+\ldots+na_n = o(n),$$

$$(\omega')\qquad \int\limits_0^t ua(u)\,du = o(t).$$

7.2. Tauber's first theorem. The first of Tauber's theorems was

THEOREM 85. *If $\sum a_n$ is summable* (A) *to sum s, and $a_n = o(n^{-1})$, then $\sum a_n$ converges to s*

or 'K_A and o imply K'. The integral analogue is 'K'_A and o' imply K''. We call this Theorem 85 *a*, and use this notation for integral analogues generally.

We begin by proving Theorem 85 *a*, and deduce Theorem 85 as a corollary: we might also prove Theorem 85 directly by an argument running parallel to that used in the proof of Theorem 85 *a*.

It is plain that (o') implies the absolute convergence of $J(y)$ for $y>0$. Also

$$j\left(\frac{1}{y}\right)-J(y)=\int_0^{1/y}a(t)\,dt-\int_0^\infty e^{-yt}a(t)\,dt$$

$$=\int_0^{1/y}(1-e^{-yt})a(t)\,dt-\int_{1/y}^\infty e^{-yt}a(t)\,dt=P-Q;$$

$$P=\int_0^{1/y}O(yt)o\left(\frac{1}{t}\right)dt=y\int_0^{1/y}o(1)\,dt=o(1),$$

since $0\leqslant 1-e^{-yt}\leqslant yt$; and

$$Q=\int_{1/y}^\infty e^{-yt}o\left(\frac{1}{t}\right)dt=o\left(y\int_{1/y}^\infty e^{-yt}\,dt\right)=o\left(\int_1^\infty e^{-u}\,du\right)=o(1).$$

Hence $j(1/y)=J(y)+o(1)\to j$ when $y\to 0$, i.e. $j(t)\to j$ when $t\to\infty$.

To deduce Theorem 85 we take $a(t)=a_n$ for $n\leqslant t<n+1$. Then

$$J(y)=\sum a_n\int_n^{n+1}e^{-yt}\,dt=\frac{1}{y}\sum a_n\{e^{-ny}-e^{-(n+1)y}\}$$

$$=\frac{1-e^{-y}}{y}\sum a_n e^{-ny}=\frac{1-e^{-y}}{y}S(y),$$

so that $S(y)\to s$ implies $J(y)\to s$. Also $a_n=o(n^{-1})$ implies $a(t)=o(t^{-1})$; so that Theorem 85 follows from Theorem 85 *a*.

7.3. Tauber's second theorem. In Tauber's second theorem the hypothesis o is replaced by ω. This changes the character of the theorem; for the convergence of S *implies* ω, by Theorem 26, so that ω is a *necessary* condition for K.

THEOREM 86. *If* $\sum a_n$ *is summable* (A) *to* s, *then* ω *is a necessary and sufficient condition for its convergence to* s.

The integral analogue is '*if* K'_A *is true, then* ω' *is necessary and sufficient for* K''. It will be convenient, here and later, to prove the main theorem and its integral analogue together, as special cases of a theorem

concerning Stieltjes integrals. We take for granted the definition and elementary properties of the 'Riemann-Stieltjes' integral

$$\int_a^T f(t)\,d\alpha(t),$$

where (a, T) is finite. In particular we assume that the integral exists when one of the functions is continuous and the other of bounded variation, and that

$$\int_a^T f(t)\,d\alpha(t) = f(T)\alpha(T)-f(a)\alpha(a)-\int_a^T \alpha(t)\,df(t). \tag{7.3.1}$$

We shall always suppose that $\alpha(t)$ is of bounded variation and that $\alpha(0) = 0$. We shall also use the equation

$$\int_a^T f(t)\,d\alpha(t) = \int_a^T \frac{f(t)}{g(t)}\,d\beta(t), \tag{7.3.2}$$

where $\beta(t) = \int_a^t g(u)\,d\alpha(u)$, f and g are continuous, and $g > 0$.

We define the Stieltjes integral over (a,∞) by

$$\int_a^\infty f(t)\,d\alpha(t) = \lim_{T\to\infty}\int_a^T f(t)\,d\alpha(t).$$

The integrals with which we shall be concerned are of the type

$$I(y) = \int_0^\infty e^{-yt}\,d\alpha(t). \tag{7.3.3}$$

We shall always suppose $I(y)$ convergent for all positive y, in which case $\alpha(t) = o(e^{yt})$ for all such y. If $\alpha(t)$ is absolutely continuous, and $\alpha'(t) = a(t)$, then $I(y)$ reduces to $J(y)$. If $\alpha(t)$ is the step-function with jumps a_n at the points $t = n$,† then it reduces to $S(y)$. Thus any Abelian or Tauberian theorem concerning $I(y)$ will contain one for $J(y)$ and one for $S(y)$.

THEOREM 87. *If $\alpha(t) \to l$ when $t \to \infty$, then $I(y)$ is convergent for $y > 0$, and $I(y) \to l$ when $y \to 0$.*

For

$$I(y) = \lim_{T\to\infty}\left\{\int_0^T e^{-yt}\,d\alpha(t)\right\}$$
$$= \lim_{T\to\infty}\left\{e^{-yT}\alpha(T)+y\int_0^T e^{-yt}\alpha(t)\,dt\right\} = y\int_0^\infty e^{-yt}\alpha(t)\,dt,$$

and so
$$I(y) \to \lim_{y\to 0} ly\int_0^\infty e^{-yt}\,dt = l.$$

† And $\alpha(+0)-\alpha(0) = \alpha(+0) = a_0$.

Next, we prove a theorem which includes Theorems 85 and 85*a*.

THEOREM 88. *If $I(y)$ is convergent for $y > 0$, and $I(y) \to l$ when $y \to 0$, then a necessary and sufficient condition that $\int d\alpha(t) = l$, i.e. that $\alpha(t) \to l$ when $t \to \infty$, is that*

(7.3.4) $$\beta(t) = \int\limits_0^t u\, d\alpha(u) = o(t)$$

when $t \to \infty$.

First, the condition is *necessary* because

$$\beta(t) = t\alpha(t) - \int\limits_0^t \alpha(u)\, du,$$

$$\frac{\beta(t)}{t} = \alpha(t) - \frac{1}{t}\int\limits_0^t \alpha(u)\, du \to l - l = 0$$

if $\alpha(t) \to l$.

Secondly, if (7.3.4) is satisfied, then (7.3.2) and (7.3.4) give

$$\alpha(t) - \alpha(1) = \int\limits_1^t d\alpha(u) = \int\limits_1^t \frac{d\beta(u)}{u} = \frac{\beta(t)}{t} - \beta(1) + \int\limits_1^t \frac{\beta(u)}{u^2}\, du$$

$$= o(1) + O(1) + o\left(\int\limits_1^t \frac{du}{u}\right) = o(\log t) = o(t)$$

and so $$\gamma(t) = \int\limits_0^t (u+1)\, d\alpha(u) = \beta(t) + \alpha(t) = o(t).$$

Now

$$\int e^{-yt}\, d\alpha(t) = \int \frac{e^{-yt}}{t+1}\, d\gamma(t) = y\int \frac{\gamma(t)}{t+1} e^{-yt}\, dt + \int \frac{\gamma(t)}{(t+1)^2} e^{-yt}\, dt.$$

The first term on the right is $o(y\int e^{-yt}\, dt) = o(1)$, and so

$$\int \delta(t) e^{-yt}\, dt = \int \frac{\gamma(t)}{(t+1)^2} e^{-yt}\, dt \to l.$$

But $\delta(t) = o(t^{-1})$, and therefore, by Theorem 85*a*, $\int \delta(t)\, dt$ converges to l. Finally,

$$\int d\alpha(t) = \int \frac{d\gamma(t)}{t+1} = \int \frac{\gamma(t)}{(t+1)^2}\, dt = \int \delta(t)\, dt = l.$$

Thus (7.3.4) is a sufficient as well as a necessary condition. Specializing $\alpha(t)$ as stated, we obtain Theorems 86 and 86*a*.

7.4. Applications to general Dirichlet's series. (1) If $\lambda_0 \geqslant 0$, $\lambda_{n+1} > \lambda_n$, $\lambda_n \to \infty$, and $\alpha(t)$ is a step-function with jumps a_n at the points λ_n, then

(7.4.1) $$I(y) = \sum a_n e^{-y\lambda_n};$$

and this specialization of $\alpha(t)$, in any of our theorems, leads to a theorem about such a Dirichlet's series. Thus Theorem 87 leads to the regularity theorem for the (A, λ) method of summation. We shall not consider the properties of the general series (7.4.1) in any detail here, but we illustrate our remarks by proving the Tauberian theorem for Dirichlet's series which corresponds to Theorem 85.

THEOREM 89. *If $S(y) = \sum a_n e^{-\lambda_n y}$ is convergent for $y > 0$; $S(y) \to s$ when $y \to 0$; and*

(7.4.2) $$a_n = o\left(\frac{\lambda_n - \lambda_{n-1}}{\lambda_n}\right);$$

then $\sum a_n$ converges to s.

We apply Theorem 88, taking

$$\alpha(t) = \sum_{\lambda_n < t} a_n.$$

Then $I(y) = S(y) \to s$. Also, if λ_ν is the last λ_n less than t, then

$$\beta(t) = \int_0^t u\, d\alpha(u) = \sum_{\lambda_n < t} \lambda_n a_n$$
$$= \lambda_0 a_0 + \sum_1^\nu o(\lambda_n - \lambda_{n-1}) = o(\lambda_\nu) = o(t).$$

Thus the conditions of Theorem 88 are satisfied, and $\alpha(t) \to s$, i.e. $\sum a_n = s$.

(2) The condition (7.4.2) is, roughly, the stronger the more slowly λ_n tends to infinity: thus it is $a_n = o(n^{-1})$ when $\lambda_n = n$, and $a_n = o\{(n \log n)^{-1}\}$ when $\lambda_n = \log n$. A divergent series which satisfies the first condition cannot, after Theorem 85, be summable (A), but it may well be summable $(\mathrm{A}, \log n)$. The latter method is not, in the language of §§ 3.8 and 4.12, so 'powerful' as the A method, since it can apply only to series $\sum a_n$ such that $\sum n^{-y} a_n$ is convergent for all positive y. Thus it is not applicable to such a series as $1-2+3-\ldots$; but, as is shown by Theorem 28 of § 4.8, it is at least as effective within its limits of applicability; and the example of the series $\sum n^{-1-ic}$ shows that it is sometimes more so (§ 7.9).

7.5. The deeper Tauberian theorems. We pass now to a series of theorems of a more difficult character, of which the best-known and in some ways the most typical is

THEOREM 90. *If* $\sum a_n$ *is summable* (A) *to sum* s, *and* $a_n = O(n^{-1})$, *then* $\sum a_n$ *converges to* s.

That is to say, 'K_A and O imply K'. This theorem is a direct generalization of Tauber's theorem 85, the condition o being replaced by O. Some prefatory remarks will be useful.

(1) If $\sum a_n$ is summable (C, k), for any k, then, by Theorem 55, it is summable (A). Hence Theorem 90 includes Theorem 63 of Ch. VI. Generally, any Tauberian theorem for A summability includes one for (C, k) summability, though an independent proof of the latter is usually easier.

(2) There are naturally variants of the theorems of §§ 7.2–4 in which 'o' is replaced by 'O' *in both hypotheses and conclusion*, and the proofs of these are trivial variants of those of the 'o' theorems. Thus Theorem 85 has the variant

'*if* $s_n = O(1)$ (A), *i.e. if* $S(y)$ *is bounded when* $y \to 0$, *and* $a_n = O(n^{-1})$, *then* $s_n = O(1)$',

and the proof, being a slightly simpler variant of that of Theorem 85, need not be set out in detail. We shall sometimes *use* such theorems, and shall indicate them by an $[O]$, Theorem 85 $[O]$, for example, being the theorem just stated; but we shall take the proofs for granted. The significant theorems of the next sections will be those in which, as in Theorem 90, 'O' occurs in one of the hypotheses but 'o' in the conclusion.

Similarly an integral analogue, Theorem Xa, will have an 'O' form, Theorem Xa $[O]$.

(3) We shall sometimes use one-sided order conditions of the types $a_n > -H\phi(n)$ or $a_n < H\phi(n)$, where a_n is real and H and $\phi(n)$ are positive. We shall write these as $a_n = O_L\{\phi(n)\}$ or $a_n = O_R\{\phi(n)\}$. Thus $a_n > -Hn^{-1}$ or $a_n = O_L(n^{-1})$ is the condition O_L of § 7.1. Actually, only O_L will occur in our theorems, since a theorem with an O_R may be deduced by a change of sign from the corresponding theorem with O_L.

We now state a series of theorems which we shall consider together with Theorem 90.

THEOREM 91. *If* $\sum a_n = s\,(\mathrm{A})$, a_n *is real, and* $a_n = O_L(n^{-1})$, *then* $\sum a_n = s$.

THEOREM 92. *If* $\sum a_n = s\,(\mathrm{A})$, *and* $s_n = O(1)$, *then* $\sum a_n = s\,(\mathrm{C}, 1)$.

THEOREM 93. *If* $\sum a_n = s\,(\mathrm{A})$ *and* $s_n \geqslant 0$, *then* $\sum a_n = s\,(\mathrm{C}, 1)$.

THEOREM 94. *If* $\sum a_n = s\,(\mathrm{A})$, a_n *is real, and* $s_n = O_L(1)$, *then* $\sum a_n = s\,(\mathrm{C}, 1)$.

THEOREM 95. *If*

$$f(x) = \sum a_n x^n \sim \frac{C}{1-x} \tag{7.5.1}$$

when $x \to 1$, and $a_n = O(1)$, then $s_n \sim Cn$.

THEOREM 96. *If (7.5.1) is true, and $a_n \geqslant 0$, then $s_n \sim Cn$.*

THEOREM 97. *If (7.5.1) is true, a_n is real, and $a_n = O_L(1)$, then $s_n \sim Cn$.*

In the last three theorems (7.5.1) is to be interpreted as $(1-x)f(x) \to 0$, and $s_n \sim Cn$ as $s_n = o(n)$, when $C = 0$. Since $1-e^{-y} \sim y$ when $y \to 0$ (7.5.1) is equivalent to $S(y) \sim Cy^{-1}$.

All these theorems are of the same depth, and it is comparatively easy to deduce any one of them from any other: the most interesting of these deductions will be found in §§ 7.7 and 7.8. In some cases the deductions are quite trivial. Thus Theorem 93 is obviously a special case of Theorem 94, and Theorem 96 of Theorem 97. Theorems 90, 92, and 95 are special cases of Theorems 91, 94, and 97 respectively when a_n is real, and may be reduced to special cases of them in any case by considering real and imaginary parts separately. Thus we have only to prove Theorems 91, 94, and 97.

Next, Theorem 94 is a corollary of Theorem 97. For if the conditions of Theorem 94 are satisfied, then

$$\sum s_n x^n = \frac{1}{1-x} \sum a_n x^n \sim \frac{s}{1-x}$$

and $s_n = O_L(1)$. Hence, assuming the truth of Theorem 97, and applying it to $\sum s_n x^n$, we obtain

$$s_0+s_1+\ldots+s_n \sim sn$$

or $\sum a_n = s$ (C, 1).

Finally, while Theorem 96 is a special case of Theorem 97, the latter is a corollary of the former. For if the conditions of Theorem 97 are satisfied, and $a_n > -H$, say, then $b_n = a_n+H > 0$ and

$$\sum b_n x^n = \sum a_n x^n + \frac{H}{1-x} \sim \frac{C+H}{1-x}.$$

Hence (assuming Theorem 96) we have

$$b_0+b_1+\ldots+b_n \sim (C+H)n,$$

and therefore $s_n \sim Cn$.

Thus it is enough to prove Theorems 91 and 96. The set of integral analogues of the theorems may be reduced in the same way. Actually, we shall prove Theorems 96 and 96*a* directly and deduce Theorems 91 and 91*a* from them.

7.6. Proof of Theorems 96 and 96 *a*. We shall prove Theorems 96 and 96 *a* as special cases of a theorem concerning Stieltjes integrals.

THEOREM 98. *If $\alpha(t)$ increases with t,*

$$I(y) = \int e^{-yt}\,d\alpha(t)$$

is convergent for $y > 0$, and $I(y) \sim Cy^{-1}$, where $C \geqslant 0$, when $y \to 0$, then $\alpha(t) \sim Ct$.

We need two lemmas.

THEOREM 99. *If $g(x)$ is real, and Riemann integrable in $(0,1)$, then there are polynomials $p(x)$ and $P(x)$ such that $p(x) < g(x) < P(x)$ and*

$$\int_0^1 \{P(x)-p(x)\}\,dx = \int_0^\infty e^{-t}\{P(e^{-t})-p(e^{-t})\}\,dt < \epsilon.$$

(i) Suppose first that g is 1 in (α,β), where $0 \leqslant \alpha < \beta \leqslant 1$, and 0 outside (α,β). We can plainly find a continuous h† such that

$$g \leqslant h, \qquad \int (h-g)\,dx < \epsilon.$$

By Weierstrass's theorem, there is a polynomial Q such that $|h-Q| < \epsilon$. If $P = Q+\epsilon$, then $g \leqslant h < P$ and

$$\int (P-g)\,dx \leqslant \int (P-Q)\,dx + \int |Q-h|\,dx + \int (h-g)\,dx < 3\epsilon.$$

Similarly there is a p such that $p < g$ and $\int (g-p)\,dx < 3\epsilon$; and p and P satisfy the requirements of the theorem (with 6ϵ for ϵ). Thus the theorem is true for this special g.

(ii) It follows by multiplication and addition that the theorem is true for any finite step-function.

(iii) If g is any Riemann integrable function, then there are finite step-functions g_1 and g_2 such that

$$g_1 \leqslant g \leqslant g_2, \qquad \int (g_2-g_1)\,dx < \epsilon.$$

We associate polynomials p_1, P_1 with g_1, and p_2, P_2 with g_2, in the manner prescribed by the theorem. Then $p_1 < g < P_2$,

$$\int (P_2-g_2)\,dx < \epsilon, \qquad \int (g_1-p_1)\,dx < \epsilon,$$

and $$\int (P_2-p_1)\,dx = \int \{(P_2-g_2)+(g_2-g_1)+(g_1-p_1)\}\,dx < 3\epsilon,$$

which proves the theorem.

† Which may be g in (α, β). In what follows integrals with respect to x, without limits shown, are over $(0, 1)$, and those with respect to t over $(0, \infty)$.

In our second lemma we require a little more knowledge of the Stieltjes integral than we have presupposed up to the present. We assume that, if f and g are of bounded variation, but not necessarily continuous, in a finite interval, and have no common points of discontinuity, then each is integrable with respect to the other. The formula for partial integration is still valid, but we shall not need it. Integrals up to ∞ are defined as limits, as in § 7.3.

THEOREM 100. *Suppose that $\alpha(t)$ increases with t, that $I(y)$ is convergent for $y > 0$, that $I(y) \sim Cy^{-1}$, and that $g(x)$ is of bounded variation in* $(0, 1)$. *Then*

$$\chi(y) = \int e^{-yt} g(e^{-yt})\, d\alpha(t)$$

exists for all positive values of y except values τ/ω, where ω is a discontinuity of α and τ a discontinuity of $g(e^{-t})$; and

$$\chi(y) \sim \frac{C}{y} \int e^{-t} g(e^{-t})\, dt \tag{7.6.1}$$

when $y \to 0$ through any sequence of positive values which excludes these exceptional values.

The values of y excluded are those for which $g(e^{-yt})$ and $\alpha(t)$ have common discontinuities: $\chi(y)$ is not defined for such y. Since the ω and the τ are at most enumerable, we exclude at most an enumerable set of values y_k of y.

Since a function of bounded variation is Riemann integrable, we can, by Theorem 99, choose polynomials p and P so that

$$p < g < P, \qquad \int e^{-t}\{P(e^{-t}) - p(e^{-t})\}\, dt < \epsilon.$$

Then
$$\int e^{-t} p(e^{-t})\, dt < \int e^{-t} g(e^{-t})\, dt < \int e^{-t} P(e^{-t})\, dt;$$

and
$$\int e^{-yt} p(e^{-yt})\, d\alpha(t) \leqslant \int e^{-yt} g(e^{-yt})\, d\alpha(t) \leqslant \int e^{-yt} P(e^{-yt})\, d\alpha(t),$$

for $y \neq y_k$, because $\alpha(t)$ increases with t. Now

$$\int e^{-yt} e^{-nyt}\, d\alpha(t) = \int e^{-(n+1)yt}\, d\alpha(t) \sim \frac{C}{(n+1)y} = \frac{C}{y} \int e^{-t} e^{-nt}\, dt,$$

and therefore

$$\int e^{-yt} P(e^{-yt})\, d\alpha(t) \sim \frac{C}{y} \int e^{-t} P(e^{-t})\, dt.$$

Hence, if $y \to 0$ through any sequence free from values y_k, we have

$$\varlimsup_{y \to 0} y \int e^{-yt} g(e^{-yt})\, d\alpha(t) \leqslant \varlimsup_{y \to 0} y \int e^{-yt} P(e^{-yt})\, d\alpha(t) \tag{7.6.2}$$

$$= C \int e^{-t} P(e^{-t})\, dt < C \int e^{-t} g(e^{-t})\, dt + C\epsilon.$$

Similarly, using p instead of P, we obtain

$$\varliminf_{y\to 0} y\int e^{-yt}g(e^{-yt})\,d\alpha(t) > C\int e^{-t}g(e^{-t})\,dt - C\epsilon; \tag{7.6.3}$$

and (7.6.1) follows from (7.6.2) and (7.6.3).

We can now prove Theorem 98. We suppose, as always, that $\alpha(0) = 0$. We take

$$g(x) = x^{-1} \quad (e^{-1} \leqslant x \leqslant 1), \qquad g(x) = 0 \quad (0 \leqslant x < e^{-1}),$$

so that $g(e^{-t}) = e^t$ for $0 \leqslant t \leqslant 1$ and $g(e^{-t}) = 0$ for $t > 1$. Then

$$\chi(y) = \int_0^\infty e^{-yt}g(e^{-yt})\,d\alpha(t) = \int_0^{1/y} d\alpha(t) = \alpha\left(\frac{1}{y}\right).$$

Also
$$\int_0^\infty e^{-t}g(e^{-t})\,dt = \int_0^1 dt = 1.$$

Hence, by Theorem 100, $\alpha(y^{-1}) \sim Cy^{-1}$ when $y \to 0$, i.e. $\alpha(t) \sim Ct$ when $t \to \infty$, exception being made in either case of a certain enumerable set of values. Here there is just one τ, viz. 1, and the values of t excluded are the discontinuities of $\alpha(t)$. Thus $\alpha(t) \sim Ct$ when $t \to \infty$ through points of continuity of $\alpha(t)$. Finally, since $\alpha(t)$ increases with t, it is true without reservation.

Theorem 98 includes Theorem 96 and its integral analogue Theorem 96 *a*. If $\alpha(t)$ is a step-function with jumps $a_n \geqslant 0$ for $t = n$, then

$$I(y) = S(y) = \sum a_n e^{-ny},$$

and $S(y) \sim Cy^{-1}$ implies $s_n \sim Cn$. This is Theorem 96. Similarly, if $\alpha(t)$ is absolutely continuous, and $\alpha'(t) = a(t)$, we obtain Theorem 96 *a*.

7.7. Proof of Theorems 91 and 91 *a*. We can now prove a theorem which includes Theorems 91 and 91 *a*. We require two further lemmas.

THEOREM 101. *If $f(y)$ is twice differentiable for positive y, and*

(7.7.1) $\quad f(y) \to l,$ $\qquad$ (7.7.2) $\quad f''(y) > -Ky^{-2},$

when $y \to 0$, then $yf'(y) \to 0$.

The theorem is one of an important type, and it will be instructive to give two proofs.

(1) If y and $y+\eta$ are positive, then

$$f(y+\eta)-f(y) = \eta f'(y)+\tfrac{1}{2}\eta^2 f''(y+\theta\eta), \tag{7.7.3}$$

where $0 < \theta < 1$; or

$$f'(y) = \frac{f(y+\eta)-f(y)}{\eta} - \tfrac{1}{2}\eta f''(y+\theta\eta). \tag{7.7.4}$$

We choose δ so that $0 < \delta < 1$ and

$$\frac{K\delta}{2(1-\delta)^2} < \epsilon, \tag{7.7.5}$$

and apply (7.7.4) with $\eta = \delta y$ and $\eta = -\delta y$.

First, taking $\eta = \delta y$, (7.7.1), (7.7.2), (7.7.4) and (7.7.5) give

$$f'(y) \leqslant o\left(\frac{1}{\delta y}\right)+\frac{K\delta y}{2(y+\theta\delta y)^2} \leqslant o\left(\frac{1}{\delta y}\right)+\frac{K\delta}{2y} \leqslant o\left(\frac{1}{\delta y}\right)+\frac{\epsilon}{y},$$

and so $$\overline{\lim}\, yf'(y) \leqslant \epsilon.$$

Secondly, taking $\eta = -\delta y$, they give

$$f'(y) \geqslant o\left(\frac{1}{\delta y}\right)-\frac{K\delta y}{2(y-\theta\delta y)^2} \geqslant o\left(\frac{1}{\delta y}\right)-\frac{K\delta}{2(1-\delta)^2 y} \geqslant o\left(\frac{1}{\delta y}\right)-\frac{\epsilon}{y},$$

and so $$\underline{\lim}\, yf'(y) \geqslant -\epsilon.$$

Hence $yf'(y) \to 0$.

(2) We observe first that if $\phi(y) = f'(y)-Ky^{-1}$ then

$$\phi'(y) = f''(y)+Ky^{-2} > 0.$$

Thus ϕ is an increasing function which has a finite derivative ϕ' for each y, and is therefore the integral of ϕ'.† Hence f' is the integral of f''.

If $yf'(y) \not\to 0$, then one or other of the inequalities

$$f'(y) > Hy^{-1}, \qquad f'(y) < -Hy^{-1} \tag{7.7.6}$$

is true for some positive H and a sequence of values of y tending to 0. Let us suppose, for example, that the first inequality (7.7.6) is satisfied for the values $y = Y$. If

$$\delta = H/2K, \qquad Y \leqslant y \leqslant Y+\delta Y,$$

then

$$f'(y) = f'(Y)+\int\limits_Y^y f''(u)\,du \geqslant \frac{H}{Y}-K\int\limits_Y^y \frac{du}{u^2}$$

$$\geqslant \frac{H}{Y}-K\int\limits_Y^{Y+\delta Y} \frac{du}{Y^2} = \frac{H}{Y}-\frac{K\delta}{Y} = \frac{H}{2Y};$$

and therefore

$$f(Y+\delta Y)-f(Y) = \int\limits_Y^{Y+\delta Y} f'(u)\,du \geqslant \frac{H}{2Y}\delta Y = \frac{H^2}{4K},$$

which contradicts (7.7.1). Similarly, considering an interval

$$Y-\delta Y \leqslant y \leqslant Y,$$

we obtain a contradiction from the second inequality (7.7.6). Hence $f'(y) = o(y^{-1})$.

† See, for example, Titchmarsh, *Theory of functions*, 368.

We can now prove

THEOREM 102. *If* (i) $I(y)$ *is convergent for* $y > 0$, *and* $I(y) \to l$ *when* $y \to 0$; (ii) *there is a function* $\beta(t)$ *such that* $\beta(0) = 0$, $\beta(t) \sim t$, *and*

(7.7.7) $$\gamma(t) = L\beta(t) + \int_0^t u\,d\alpha(u)$$

is, for some positive L, *an increasing function of* t; *then* $\alpha(t) \to l$.

It follows from the definition of $\beta(t)$ that it is the difference of two bounded and increasing functions, and therefore of bounded variation, in any finite interval $(0, T)$.

We observe first that

(7.7.8) $$\int e^{-yt}\,d\beta(t) = y\int e^{-yt}\beta(t)\,dt \sim y\int te^{-yt}\,dt = \frac{1}{y},$$

(7.7.9) $$\int e^{-yt}t\,d\beta(t) = \int e^{-yt}(yt-1)\beta(t)\,dt$$
$$\sim y\int t^2e^{-yt}\,dt - \int te^{-yt}\,dt = \frac{2}{y^2} - \frac{1}{y^2} = \frac{1}{y^2},$$
$$I'(y) = -\int te^{-yt}\,d\alpha(t), \qquad I''(y) = \int t^2e^{-yt}\,d\alpha(t).$$

Hence, first,

$$I''(y) = \int te^{-yt}\,d\gamma(t) - L\int te^{-yt}\,d\beta(t) \geqslant -L\int te^{-yt}\,d\beta(t) > -\frac{M}{y^2},$$

for an appropriate M. It follows, by Theorem 101, that $I'(y) = o(y^{-1})$. But

$$I'(y) = -\int e^{-yt}\,d\gamma(t) + L\int e^{-yt}\,d\beta(t);$$

and therefore, by (7.7.8),

(7.7.10) $$\int e^{-yt}\,d\gamma(t) \sim \frac{L}{y}.$$

Since $\gamma(t)$ increases, it follows from (7.7.10) and Theorem 98, that $\gamma(t) \sim Lt$; and so that

(7.7.11) $$\int_0^t u\,d\alpha(u) = \gamma(t) - L\beta(t) = o(t).$$

Finally, it follows from (7.7.11) and Theorem 88 that $\alpha(t) \to l$.

This proves Theorem 102. If $\alpha(t)$ is the step-function of § 7.4 (1), with $\lambda_n = n$, and $na_n > -H$, we may take

$$\beta(t) = \sum_{n<t} 1.$$

Then $$\gamma(t) = L\beta(t) + \sum_{n<t} na_n = \sum_{n<t} (na_n + L)$$

increases with t if $L > H$, and we obtain Theorem 91. If $\alpha(t)$ is

absolutely continuous, $\alpha'(t) = a(t)$, and $ta(t) > -H$, then we may take $\beta(t) = t$, and obtain Theorem 91 *a* similarly.

We make one further specialization of Theorem 102. Suppose that $\lambda_n < \lambda_{n+1}$, $\lambda_n \to \infty$ and

(7.7.12) $$\lambda_{n+1} \sim \lambda_n,$$

that $\alpha(t)$ is the step-function of § 7.4 (1), and that

(7.7.13) $$a_n > -H(\lambda_n - \lambda_{n-1})/\lambda_n.$$

If $$\beta(t) = \sum_{\lambda_n < t} (\lambda_n - \lambda_{n-1})$$

(with $\lambda_{-1} = 0$) and $L > H$, then

$$\gamma(t) = L\beta(t) + \int_0^t u\, d\alpha(u) = \sum_{\lambda_n < t} \{L(\lambda_n - \lambda_{n-1}) + \lambda_n a_n\}$$

is an increasing function of t. Also $\beta(t) \sim \lambda_N$, where N is the last value of n for which $\lambda_n < t$; and (7.7.12) then shows that $\beta(t) \sim t$. Hence the conditions of Theorem 102 are satisfied, and we obtain

THEOREM 103. *If λ_n tends to infinity so as to satisfy* (7.7.12), *a_n satisfies* (7.7.13), *and $S(y) = \sum a_n e^{-\lambda_n y} \to s$ when $y \to 0$, then $\sum a_n$ converges to s.*

This theorem corresponds to Theorem 91 as Theorem 89 corresponds to Theorem 85; but there is an additional condition on λ_n, viz. (7.7.12). This restriction is essential; the proof fails without it, since it is then no longer true that $\beta(t) \sim t$; and the theorem itself becomes false. Suppose, for example, that

$$\lambda_{2m} = 2^m + 2^{-m-2}, \qquad \lambda_{2m+1} = 2^{m+1},$$

and $a_n = (-1)^n$. Then $a_n > 0$ if n is even. Also

$$\frac{\lambda_{2m+1} - \lambda_{2m}}{\lambda_{2m+1}} = \frac{2^{m+1} - 2^m - 2^{-m-2}}{2^{m+1}} \geq 1 - \tfrac{1}{2} - \tfrac{1}{8} > \tfrac{1}{4},$$

so that $$a_{2m+1} = -1 > -4(\lambda_{2m+1} - \lambda_{2m})/\lambda_{2m+1}.$$

Thus (7.7.13) is true with $H = 4$. Also

$$S(y) = e^{-\frac{1}{4}y} - \sum e^{-2^{m+1}y}(1 - e^{-2^{-m-2}y}) = e^{-\frac{1}{4}y} + O(y \sum 2^{-m-3}) = 1 + O(y) \to 1;$$

but $\sum a_n$ is not convergent.

There is a difference in this respect between Theorem 103 and the more direct generalization of Theorem 90, viz.

THEOREM 104. *If $S(y) = \sum a_n e^{-\lambda_n y} \to s$ and $a_n = O\{(\lambda_n - \lambda_{n-1})/\lambda_n\}$, then $\sum a_n$ converges to s.*

Here it is not necessary to assume (7.7.12).†

7.8. Further remarks on the relations between the theorems of § 7.5. There are various methods of proving the theorems of § 7.5, the simplest being Karamata's, which we have followed here. The original method of Hardy and Littlewood involves a technique of repeated differentiation, about which we shall say something in § 7.12.

† See the note at the end of the chapter.

There is also the method of Wiener, which is the most powerful and the most general, but also the most difficult, since it depends on deep theorems in the theory of Fourier transforms. This we leave to Ch. XII.

Each method involves a characteristic idea, leading to one of the theorems from which the others are deduced by more elementary devices. Thus Karamata's idea is embodied in Theorem 100, and the Tauberian theorem to which it leads naturally is Theorem 96. The method of Hardy and Littlewood leads to Theorem 90 or Theorem 96, according to the manner of its use; while that of Wiener leads most naturally to Theorem 92.

It is therefore interesting to examine the relations between the theorems more closely. We show here (i) how to deduce Theorem 92 from Theorem 90, and (ii) how to deduce Theorem 96 from Theorem 92.

(i) *Deduction of Theorem* 92 *from Theorem* 90. We suppose that the conditions of Theorem 92 are satisfied and, as we may without real loss of generality, that $a_0 = 0$ and $s = 0$. We write

$$w_0 = 0, \qquad w_n = a_1+2a_2+\dots+na_n \quad (n > 0), \qquad v_n = \frac{w_n}{n(n+1)},$$

so that

$$w_n = (n+1)s_n - s_0 - s_1 - \dots - s_n = O(n), \qquad v_n = O(n^{-1});$$

and
$$f(x) = \sum a_n x^n, \qquad g(x) = \sum v_n x^{n+1}.$$

Then

$$g(x)+(1-x)g'(x) = \sum \frac{w_n}{n(n+1)} x^{n+1} + \sum \frac{w_n}{n} x^n - \sum \frac{w_n}{n} x^{n+1}$$

$$= \sum \frac{w_n}{n} x^n - \sum \frac{w_n}{n+1} x^{n+1} = \sum \frac{w_n - w_{n-1}}{n} x^n = f(x).\dagger$$

Hence

$$g(x)+(1-x)g'(x) = o(1), \qquad \frac{d}{dx}\left\{\frac{g(x)}{1-x}\right\} = o\left\{\frac{1}{(1-x)^2}\right\},$$

and therefore, integrating, we have

$$\frac{g(x)}{1-x} = o\left(\frac{1}{1-x}\right), \qquad g(x) = o(1).$$

Since $g(x) = \sum v_n x^{n+1} = o(1)$ and $v_n = O(n^{-1})$, it follows from Theorem 90 that $\sum v_n$ converges to 0. But

$$\sum_1^N v_n = \sum_1^N \left(\frac{1}{n} - \frac{1}{n+1}\right) w_n = \sum_1^N \frac{w_n - w_{n-1}}{n} - \frac{w_N}{N+1}$$

$$= s_N - \frac{w_N}{N+1} = \frac{s_0+s_1+\dots+s_N}{N+1}.$$

Hence $s_0+s_1+\dots+s_N = o(N)$, i.e. $\sum a_n = 0$ (C, 1).

† The summations running from 1 to ∞.

(ii) *Deduction of Theorem* 96 *from Theorem* 92. We suppose the hypotheses of Theorem 96 satisfied. Then, if $n > 1$, we have

$$s_n \leqslant \left(1-\frac{1}{n}\right)^{-n} \sum_0^n a_m \left(1-\frac{1}{n}\right)^m \leqslant 4f\left(1-\frac{1}{n}\right) = O(n).$$

We write $t_n = s_n/(n+1)$, so that $t_n = O(1)$. Hence, if

$$b_n = t_{n+1}-t_n = \frac{s_{n+1}}{n+2}-\frac{s_n}{n+1} = \frac{a_{n+1}}{n+2}-\frac{s_n}{(n+1)(n+2)},$$

then

$$b_0+b_1+...+b_n = t_{n+1}-s_0 = O(1), \tag{7.8.1}$$

$$b_n \geqslant -\frac{s_n}{(n+1)(n+2)} > -\frac{H}{n} \tag{7.8.2}$$

for an appropriate H. Next,

$$\sum t_n x^n = \sum \frac{s_n}{n+1} x^n = \frac{1}{x}\int_0^x (\sum s_n t^n)\, dt = \frac{1}{x}\int_0^x \frac{1}{1-t}(\sum a_n t^n)\, dt$$

$$= \frac{1}{x}\int_0^x \frac{f(t)}{1-t}\, dt \sim C\int_0^x \frac{dt}{(1-t)^2} \sim \frac{C}{1-x},$$

and so

$$\sum b_n x^n = (t_1-t_0)+(t_2-t_1)x+(t_3-t_2)x^2+... \tag{7.8.3}$$

$$= -\frac{t_0}{x}+\frac{1-x}{x}\sum t_n x^n \to C-t_0.$$

From (7.8.3), (7.8.1), and Theorem 92 it follows that $\sum b_n = C-t_0$ (C, 1), i.e.

$$t_n = t_0+\sum_0^{n-1} b_m \to C \quad (C, 1). \tag{7.8.4}$$

Finally, from (7.8.4), (7.8.2), and Theorem 64, it follows that $t_n \to C$, i.e. that $s_n \sim Cn$.

7.9. The series $\sum n^{-1-ic}$. We have seen in § 6.11 that the series $\sum n^{-1-ic}$, where c is real and not 0, is not convergent, and that in fact

$$s_n = -\frac{n^{-ic}}{ic}+l+o(1),$$

where l is independent of n.† Since $a_n = O(n^{-1})$, it follows from

† We shall identify l as $\zeta(1+ic)$ in Ch. XIII.

Theorem 90 that $\sum a_n$ is not summable (A): that it is not summable (C) followed from Theorem 63. It is summable (A, $\log n$), since

$$\sum n^{-1-y-ic} = \zeta(1+y+ic) \to \zeta(1+ic).$$

Theorems 63 and 90, and the other theorems of § 7.5, have many applications in the theory of Fourier series. It is known, for example, that the Fourier series of any integrable $f(t)$ is summable (A), or (C, 1), at any point of continuity or jump of $f(t)$. If $f(t)$ is of bounded variation, then its Fourier coefficients are $O(n^{-1})$; and it then follows, from Theorem 90 or Theorem 63, that the series is convergent for all t.

7.10. Slowly oscillating and slowly decreasing functions. We can generalize Theorem 91 further by the use of the ideas of § 6.2.

THEOREM 105. *If* (i) $I(y) = \int e^{-yt}\,d\alpha(t)$ *is convergent for* $y > 0$, *and* $I(y) \to l$ *when* $y \to 0$; (ii) $\alpha(t)$ *is slowly decreasing; then* $\alpha(t) \to l$ *when* $t \to \infty$.

THEOREM 106. *If* $\sum a_n = s$ (A), *and* s_n *is slowly decreasing, then* $\sum a_n = s$.

It is convenient to suppose, as plainly we may, that $\alpha(t) = 0$ in an interval $(0, \tau)$. We need a lemma.

THEOREM 107. *If* $\alpha(t)$ *is* 0 *in an interval* $(0,\tau)$, *and of bounded variation in any interval* $(0, T)$, *and* $I(y)$ *is convergent for* $y > 0$, *then, if* $p > q > 0$,

$$\int_0^\infty \frac{\alpha(pt)-\alpha(qt)}{t} e^{-yt}\,dt = \int_{y/p}^{y/q} \frac{I(u)}{u}\,du.$$

The integral for $I(u)$ is uniformly convergent in any interval $0 < v \leqslant u \leqslant U$, and $\alpha(t) = o(e^{\epsilon t})$ for all positive ϵ. If $0 < v < U$, then

$$\int_v^U \frac{I(u)}{u}\,du = \int_v^U \frac{du}{u} \int_0^\infty e^{-ut}\,d\alpha(t)$$

$$= \int_0^\infty d\alpha(t) \int_v^U \frac{e^{-ut}}{u}\,du = -\int_0^\infty \alpha(t)\left(\frac{d}{dt}\int_v^U \frac{e^{-ut}}{u}\,du\right)dt$$

$$= \int_0^\infty \alpha(t)\,dt \int_v^U e^{-ut}\,du = \int_0^\infty \frac{e^{-vt}-e^{-Ut}}{t}\alpha(t)\,dt.$$

Finally, taking $v = y/p$, $U = y/q$, we obtain

$$\int \frac{e^{-yt/p}}{t}\alpha(t)\,dt - \int \frac{e^{-yt/q}}{t}\alpha(t)\,dt$$

$$= \int \frac{e^{-yt}}{t}\alpha(pt)\,dt - \int \frac{e^{-yt}}{t}\alpha(qt)\,dt = \int \frac{\alpha(pt)-\alpha(qt)}{t}e^{-yt}\,dt.$$

Passing to the proof of Theorem 105, we may suppose $l = 0$. Since $\alpha(t)$ is slowly decreasing, and bounded in any finite interval of positive t, we have

$$\frac{\alpha(pt)-\alpha(qt)}{t} > -\frac{H}{t} \tag{7.10.1}$$

for any fixed p and q with $0 < q < p$ and an appropriate H.† Since $I(y) \to 0$,

$$\int_0^\infty \frac{\alpha(pt)-\alpha(qt)}{t} e^{-yt}\,dt = \int_{y/p}^{y/q} \frac{I(u)}{u}\,du = o\left(\log\frac{p}{q}\right) = o(1), \tag{7.10.2}$$

for any fixed p and q with $0 < q < p$; and from (7.10.1) and (7.10.2) it follows, by Theorem 91 a, that

$$\int \frac{\alpha(pt)-\alpha(qt)}{t}\,dt = 0.$$

Hence

$$\int_{qT}^{pT} \frac{\alpha(u)}{u}\,du = \int_0^T \frac{\alpha(pt)-\alpha(qt)}{t}\,dt \to 0 \tag{7.10.3}$$

when $T \to \infty$.

If $\alpha(t)$ does not tend to 0, there is a positive M such that one or other of the inequalities $\alpha(t) > M$, $\alpha(t) < -M$ is true for a sequence of values T tending to ∞. Let us suppose, for example, that the first inequality is true for $t = T$. We take $q = 1$, and choose $p > 1$ so that $\alpha(u)-\alpha(v) > -\frac{1}{2}M$ for $v > v_0$, $v \leqslant u \leqslant pv$. Then, for sufficiently large $t = T$, we have

$$\alpha(u) > \alpha(T)-\tfrac{1}{2}M > \tfrac{1}{2}M \quad (T \leqslant u \leqslant pT);$$

and therefore
$$\int_T^{pT} \frac{\alpha(u)}{u}\,du > \tfrac{1}{2}M \log p,$$

in contradiction to (7.10.3).

Similarly (considering an interval to the left of a T) we obtain a contradiction from $\alpha(T) < -M$, and the theorem follows. Finally we obtain Theorem 106 by supposing $\alpha(t)$ an appropriate step-function.

7.11. Another generalization of Theorem 98. We have so far proved our theorems in their simplest forms, ignoring the many generalizations which involve additional functions or parameters. We now illustrate these by an important extension of Theorem 98. We

† See § 6.2.

suppose throughout this section that $\phi(x)$ is positive and increasing for $x \geqslant x_0$, and tends to infinity with x; and that

$$\phi(x) = x^{\sigma} L(x), \tag{7.11.1}$$

where $\sigma \geqslant 0$ and

$$L(cx) \sim L(x) \tag{7.11.2}$$

for every positive c. Thus $x^{\sigma}(\log x)^{\tau}$ is a possible form of $\phi(x)$, for $x > 2$, if $\sigma > 0$, τ real, or $\sigma = 0$, $\tau > 0$.

THEOREM 108. *If $\alpha(t)$ increases with t, $I(y) = \int e^{-yt}\, d\alpha(t)$ is convergent for $y > 0$, and*

$$I(y) \sim \phi(y^{-1}) \tag{7.11.3}$$

when $y \to 0$, then

$$\alpha(t) \sim \frac{\phi(t)}{\Gamma(\sigma+1)} \tag{7.11.4}$$

when $t \to \infty$.

We suppose first that $\sigma > 0$, when the proof is a simple generalization of that of Theorem 98. We write

$$\rho(x) = \left(\log \frac{1}{x}\right)^{\sigma-1} \quad (0 < x < 1),$$

and use

THEOREM 109. *If g satisfies the conditions of Theorem 99, and $\sigma > 0$, then there are polynomials p and P such that $p < g < P$ and*

$$\int \{P(x)-p(x)\}\rho(x)\, dx = \int e^{-t}t^{\sigma-1}\{P(e^{-t})-p(e^{-t})\}\, dt < \epsilon\Gamma(\sigma).† \tag{7.11.5}$$

THEOREM 110. *If $\alpha(t)$ and $I(y)$ satisfy the conditions of Theorem 108, and $g(x)$ is of bounded variation in $(0, 1)$, then*

$$\chi(y) = \int e^{-yt} g(e^{-yt})\, d\alpha(t) \tag{7.11.6}$$

exists for all positive y except those specified in Theorem 100, and

$$\chi(y) \sim \frac{1}{\Gamma(\sigma)}\phi\left(\frac{1}{y}\right) \int e^{-t}t^{\sigma-1}g(e^{-t})\, dt \tag{7.11.7}$$

when $y \to 0$ through any sequence free from these exceptional values.

The proof of Theorem 109 is a straightforward generalization of that of Theorem 99, the changes necessitated by the presence of the weight

† As in § 7.6, integrals with respect to x, without limits, are over $(0, 1)$, those with respect to t over $(0, \infty)$.

function $\rho(x)$ being almost trivial. If g is 1 in (α, β) and 0 outside, then there is a continuous h such that

$$g \leqslant h, \qquad \int (h-g)\rho\, dx < \epsilon \int \rho\, dx = \epsilon\Gamma(\sigma),$$

and a polynomial Q such that $|Q-h| < \epsilon$. If $P = Q+\epsilon$ then $g \leqslant h < P$ and

$$\int (P-g)\rho\, dx \leqslant \int (P-Q)\rho\, dx + \int |Q-h|\rho\, dx + \int (h-g)\rho\, dx < 3\epsilon\Gamma(\sigma).$$

Similarly we can determine p so that $p < g$ and $\int (g-p)\rho\, dx < 3\epsilon\Gamma(\sigma)$, and the result (with 6ϵ for ϵ) follows. Thus the theorem is true for this g, and so for any finite step-function.

The final stage of the proof needs a little elaboration. We write $M = \max|g|$, and determine ξ and ξ' so that $0 < \xi < \xi' < 1$ and

$$2M\int_0^{\xi} \rho\, dx < \epsilon\Gamma(\sigma), \qquad 2M\int_{\xi'}^{1} \rho\, dx < \epsilon\Gamma(\sigma). \tag{7.11.8}$$

We can then find finite step-functions g_1 and g_2 such that

$$-M \leqslant g_1 \leqslant g \leqslant g_2 \leqslant M$$

in (ξ, ξ') and
$$\int_{\xi}^{\xi'} (g_2-g_1)\, dx < \frac{\epsilon\Gamma(\sigma)}{\max\{\rho(\xi), \rho(\xi')\}},$$

from which it follows that

$$\int_{\xi}^{\xi'} (g_2-g_1)\rho\, dx < \epsilon\Gamma(\sigma).\dagger \tag{7.11.9}$$

If we define g_1 as $-M$ and g_2 as M in $(0, \xi)$ and $(\xi', 1)$, then $g_1 \leqslant g \leqslant g_2$ throughout $(0, 1)$, $g_2-g_1 \leqslant 2M$, and

$$\int (g_2-g_1)\rho\, dx < 3\epsilon\Gamma(\sigma), \tag{7.11.10}$$

by (7.11.8) and (7.11.9).

Finally, since g_1 and g_2 are finite step-functions, there are polynomials p and P such that $p < g_1 \leqslant g \leqslant g_2 < P$ and

$$\int (P-g_2)\rho\, dx < \epsilon\Gamma(\sigma), \qquad \int (g_1-p)\rho\, dx < \epsilon\Gamma(\sigma).$$

It then follows from (7.11.10) that $\int (P-p)\rho\, dx < 5\epsilon\Gamma(\sigma)$, and this completes the proof of Theorem 109.

† ρ is monotonic in (0, 1) and tends to infinity at one end or the other, except when $\sigma = 1$.

Passing to Theorem 110, we have

$$\int e^{-yt}e^{-nyt}\,d\alpha(t) = \int e^{-(n+1)yt}\,d\alpha(t) \sim \phi\left\{\frac{1}{(n+1)y}\right\}$$
$$= (n+1)^{-\sigma}y^{-\sigma}L\left\{\frac{1}{(n+1)y}\right\} \sim (n+1)^{-\sigma}y^{-\sigma}L\left(\frac{1}{y}\right)$$
$$= (n+1)^{-\sigma}\phi\left(\frac{1}{y}\right) = \frac{1}{\Gamma(\sigma)}\phi\left(\frac{1}{y}\right)\int e^{-t}e^{-nt}t^{\sigma-1}\,dt,$$

for fixed n, when $y \to 0$. From this it follows that

$$\int e^{-yt}Q(e^{-yt})\,d\alpha(t) \sim \frac{1}{\Gamma(\sigma)}\phi\left(\frac{1}{y}\right)\int e^{-t}t^{\sigma-1}Q(e^{-t})\,dt \tag{7.11.11}$$

for any polynomial Q.

There are polynomials p and P such that

$$p < g < P, \qquad \int e^{-t}t^{\sigma-1}\{P(e^{-t})-p(e^{-t})\}\,dt < \epsilon\Gamma(\sigma),$$

and *a fortiori* $\int e^{-t}t^{\sigma-1}\{P(e^{-t})-g(e^{-t})\}\,dt < \epsilon\Gamma(\sigma).$

Hence, if $y \to \infty$ in the manner prescribed in Theorems 100 and 110,

$$\overline{\lim}\,\frac{1}{\phi(1/y)}\int e^{-yt}g(e^{-yt})\,d\alpha(t) \leqslant \lim\frac{1}{\phi(1/y)}\int e^{-yt}P(e^{-yt})\,d\alpha(t)$$
$$= \frac{1}{\Gamma(\sigma)}\int e^{-t}t^{\sigma-1}P(e^{-t})\,dt < \frac{1}{\Gamma(\sigma)}\int e^{-t}t^{\sigma-1}g(e^{-t})\,dt + \epsilon.$$

Similarly

$$\underline{\lim}\,\frac{1}{\phi(1/y)}\int e^{-yt}g(e^{-yt})\,d\alpha(t) > \frac{1}{\Gamma(\sigma)}\int e^{-t}t^{\sigma-1}g(e^{-t})\,dt - \epsilon,$$

and these two inequalities prove (7.11.7).

We can now prove Theorem 108 (when $\sigma > 0$). Choosing g as in the proof of Theorem 98, we obtain

$$\int_0^\infty e^{-yt}g(e^{-yt})\,d\alpha(t) = \int_0^{1/y} d\alpha(t) = \alpha\left(\frac{1}{y}\right),$$

$$\frac{1}{\Gamma(\sigma)}\int_0^\infty e^{-t}t^{\sigma-1}g(e^{-t})\,dt = \frac{1}{\Gamma(\sigma)}\int_0^1 t^{\sigma-1}\,dt = \frac{1}{\Gamma(\sigma+1)};$$

and the theorem follows from Theorem 110, since $\alpha(t)$ increases with t.

The argument fails when $\sigma = 0$ and $\phi(y^{-1}) = L(y^{-1}) \to \infty$. In this case we replace Theorem 110 by

THEOREM 111. *If $\alpha(t)$ increases with t,*

$$I(y) = \int e^{-yt}\, d\alpha(t) \sim L\left(\frac{1}{y}\right),$$

and $g(x)$ is continuous in $(0, 1)$, then

$$\chi(y) = \int e^{-yt} g(e^{-yt})\, d\alpha(t) \sim L\left(\frac{1}{y}\right) g(1).$$

Here $$\int e^{-yt} e^{-nyt}\, d\alpha(t) \sim L\left\{\frac{1}{(n+1)y}\right\} \sim L\left(\frac{1}{y}\right),$$

and so $$\int e^{-yt} Q(e^{-yt})\, d\alpha(t) \sim L\left(\frac{1}{y}\right) Q(1)$$

for any polynomial Q. Since g is continuous, there are polynomials p, P such that $p < g < P < p+\epsilon$ for $0 \leqslant x \leqslant 1$. Then

$$\int e^{-yt} g(e^{-yt})\, d\alpha(t) \leqslant \int e^{-yt} P(e^{-yt})\, d\alpha(t),$$

$$\overline{\lim} \frac{1}{L(1/y)} \int e^{-yt} g(e^{-yt})\, d\alpha(t)$$

$$\leqslant \lim \frac{1}{L(1/y)} \int e^{-yt} P(e^{-yt})\, d\alpha(t) = P(1) < g(1)+\epsilon,$$

and similarly $$\underline{\lim} \frac{1}{L(1/y)} \int e^{-yt} g(e^{-yt})\, d\alpha(t) > g(1)-\epsilon.$$

This proves Theorem 111. We pass to the proof of Theorem 108, with $\sigma = 0$. We cannot now choose g as in the proof of Theorem 98, that g being discontinuous. We take

$$g(x) = \frac{1}{x}\left(1 - \log\frac{1}{x}\right) \quad (e^{-1} \leqslant x \leqslant 1), \qquad 0 \quad (0 < x \leqslant e^{-1}),$$

so that $e^{-t} g(e^{-t})$ is $1-t$ for $0 \leqslant t \leqslant 1$ and 0 for $t > 1$. This g is continuous, so that, by Theorem 111,

$$y \int_0^{1/y} \alpha(t)\, dt = \int_0^{1/y} (1-yt)\, d\alpha(t) \sim L\left(\frac{1}{y}\right),$$

$$\int_0^x \alpha(t)\, dt \sim xL(x). \tag{7.11.12}$$

It follows from (7.11.12) that

$$\int_0^{x+\delta x} \alpha(t)\,dt \sim (x+\delta x)L(x+\delta x) \sim (x+\delta x)L(x),$$

$$\int_x^{x+\delta x} \alpha(t)\,dt \sim \delta x L(x), \qquad \int_x^{x+\delta x} \alpha(t)\,dt = \delta x L(x)+o\{xL(x)\}$$

if $\delta > 0$. Since $\alpha(t)$ increases with t

$$\delta x\alpha(x) \leqslant \delta x L(x)+o\{xL(x)\},$$

(7.11.13) $$\overline{\lim}\,\frac{\alpha(x)}{L(x)} \leqslant 1.$$

Similarly $$\int_x^{x+\delta x} \alpha(t)\,dt \sim \delta x L(x+\delta x)$$

$$\delta x\alpha(x+\delta x) \geqslant \delta x L(x+\delta x)+o\{xL(x+\delta x)\},$$

(7.11.14) $$\underline{\lim}\,\frac{\alpha(x+\delta x)}{L(x+\delta x)} \geqslant 1.$$

Finally, (7.11.13) and (7.11.14) show that $\alpha(x) \sim L(x)$.

7.12. The method of Hardy and Littlewood. We insert here a short sketch of the method by which Hardy and Littlewood first proved Theorem 96. The method is less simple than Karamata's, which we followed in § 7.6, but depends on ideas which are interesting in themselves. We begin by proving

THEOREM 112. *If $g(x)$ is differentiable for $0 \leqslant x < 1$, $g(x) \sim C(1-x)^{-\alpha}$, where $C > 0$, $\alpha > 0$, when $x \to 1$, and $g'(x)$ increases with x, then*

$$g'(x) \sim C\alpha(1-x)^{-\alpha-1}.$$

If $x = 1-y$, $g(x) = G(y)$, then $G(y) \sim Cy^{-\alpha}$ and $-G'(y)$ increases as y decreases. We choose a positive δ such that

$$(1-\epsilon)\delta\alpha < 1-(1+\delta)^{-\alpha} < (1+\epsilon)\,\delta\alpha.$$

Then $$G(y)-G(y+\delta y) \sim C\{1-(1+\delta)^{-\alpha}\}y^{-\alpha},$$

and therefore

$$G(y)-G(y+\delta y) > C(1-\epsilon)\{1-(1+\delta)^{-\alpha}\}y^{-\alpha} > C(1-\epsilon)^2\alpha\delta y^{-\alpha}$$

for sufficiently small y. But $-G'(y)$ increases as y decreases, and therefore

$$-\delta yG'(y) > \int_y^{y+\delta y} \{-G'(t)\}\,dt = G(y)-G(y+\delta y).$$

Hence $$-\delta y G'(y) > C(1-\epsilon)^2 \alpha \delta y^{-\alpha}$$

for sufficiently small y, and

$$\underline{\lim}\{-y^{\alpha+1}G'(y)\} \geqslant C\alpha(1-\epsilon)^2.$$

Similarly the upper limit does not exceed $C\alpha(1+\epsilon)^2$, and the theorem follows.

A simple corollary is

THEOREM 113. *If* $c_n \geqslant 0$ *and*

$$g(x) = \sum c_n x^n \sim C(1-x)^{-\alpha} \quad (C > 0,\ \alpha > 0),$$

then $$g^{(p)}(x) \sim C\alpha(\alpha+1)\ldots(\alpha+p-1)(1-x)^{-\alpha-p},$$

for every positive integral p.

Plainly $g'(x)$ increases with x, so that $g'(x) \sim C\alpha(1-x)^{-\alpha-1}$; and the argument may be repeated.

From this point on we do no more than indicate the main lines of the proof. One preliminary remark will help to make it more readily intelligible. It follows from (7.5.1), by the simple argument used at the beginning of § 7.8(ii), that $s_n = O(n)$; but the argument fails us as soon as we try to obtain a more precise result. The reason is, at bottom, that there is no such 'peak' in the sequence (x^n) or (e^{-ny}) as would enable us to infer that the series is dominated by terms near a maximum term. We can, however, create such a peak artificially by p differentiations with respect to y. This replaces e^{-ny} (apart from sign) by $n^p e^{-ny}$, which has a maximum where n is about $N = p/y$. The maximum is about $(p/ey)^p$, which increases rapidly with p, so that the peak is pronounced when p is large. Thus the fundamental idea of our proof will be that of *differentiating a large number of times.*

Coming more to detail, we take $C = 1$, so that

$$\sum s_n e^{-ny} = \sum s_n x^n = \frac{1}{1-x}\sum a_n x^n \sim \frac{1}{(1-x)^2} \sim \frac{1}{y^2};$$

and, after Theorem 113, we may differentiate this relation any number of times with respect to y. We thus obtain

(7.12.1) $$\sum n^p s_n e^{-ny} \sim (p+1)!\, y^{-p-2}$$

for every p. Now

(7.12.2) $$\sum n^p e^{-ny} \sim p!\, y^{-p-1}.$$

The terms of this series have a peak about where $n = N$, and decrease fairly rapidly on either side of it. It is therefore natural to suppose

(and is easily verified) that we can choose first a large p, and then an M depending on both p and y, so that $M = o(N)$ and

$$\Big(\sum_{n<N-M} + \sum_{n>N+M}\Big) n^p e^{-ny} < \epsilon p!\, y^{-p-1} \tag{7.12.3}$$

for small y. Also $s_n = O(n)$, so that s_n cannot behave violently, and it is therefore natural to suppose also that we can make a similar reduction of the series (7.12.1).

It will follow that we can choose, first $p = p(\epsilon)$ and then

$$y_0 = y_0(p,\epsilon) = y_0(\epsilon),$$

so that

$$(1-\epsilon)\sum_{N-M}^{N+M} n^p s_n e^{-ny} < (p+1)!\, y^{-p-2} < (1+\epsilon)\sum_{N-M}^{N+M} n^p s_n e^{-ny}$$

for $y \leqslant y_0(\epsilon)$. *A fortiori*, since s_n increases with n, we shall have

$$(1-\epsilon)s_{N-M}\sum_{N-M}^{N+M} n^p e^{-ny} < (p+1)!\, y^{-p-2} < (1+\epsilon)s_{N+M}\sum_{N-M}^{N+M} n^p e^{-ny}.$$

It will then follow from (7.12.2) and (7.12.3) that

$$(1-2\epsilon)s_{N-M} < (p+1)y^{-1} < (1+2\epsilon)s_{N+M}$$

for large enough p and small enough y. Finally, since

$$N \pm M \sim N \sim py^{-1},$$

it will follow that $s_N \sim N$.

There is a good deal of detail to be added, but it is mostly a matter of routine; and the proof, though admittedly less simple than Karamata's, should not be found difficult when once the ideas underlying it have been understood.

7.13. The 'high indices' theorem. If $\lambda_{n+1}/\lambda_n \to 1$ or, what is the same thing, if

$$\mu_n = \frac{\lambda_n - \lambda_{n-1}}{\lambda_n} \to 0, \tag{7.13.1}$$

$$a_n = O(\mu_n), \tag{7.13.2}$$

and $S(y) = \sum a_n e^{-\lambda_n y} \to s$, then $\sum a_n = s$. This is a special case of Theorem 103, and we stated in Theorem 104 that the result is true without the restriction (7.13.1).

A particularly interesting case is that in which λ_n increases sufficiently rapidly and regularly to make

$$\lambda_{n+1} > c\lambda_n, \tag{7.13.3}$$

where $c > 1$ (as, for example, when $\lambda_n = 2^n$). Then μ_n lies between $(c-1)/c$ and 1, so that (7.13.2) reduces to $a_n = O(1)$. Thus in this case

the theorem asserts that the series is convergent whenever its terms are bounded. This assertion, however, does not contain the full truth, which is that, when λ_n satisfies (7.13.3), then *no* restriction on a_n is necessary.

THEOREM 114. *If λ_n satisfies* (7.13.3), *and $S(y) \to s$, then $\sum a_n$ converges to s.*

We may suppose $\lambda_0 > 0$. The kernel of the proof lies in that of the lemma which follows.

THEOREM 115. *If λ_n satisfies* (7.13.3), $\lambda_0 > 0$,

$$f(y) = f_N(y) = \sum_{n=0}^{N} a_n e^{-\lambda_n y}$$

and $|f(y)| \leqslant H$ for $y > 0$, then

$$|a_n| \leqslant CH,$$

where $C = C(c)$ depends only on c.

Suppose that
$$P(y) = \sum_{r=0}^{R} p_r e^{-\nu_r y},$$
where ν_r is positive and increases with r. Then

$$F(y) = \sum_{n=0}^{N} a_n P(\lambda_n y) = \sum_{n=0}^{N} a_n \sum_{r=0}^{R} p_r e^{-\nu_r \lambda_n y} = \sum_{r=0}^{R} p_r f(\nu_r y),$$

and therefore

$$|F(y)| \leqslant H \sum_{r=0}^{R} |p_r|. \tag{7.13.4}$$

We take, in particular,

$$P(y) = \{p(y)\}^R = \{4(2^{-y} - 2^{-2y})\}^R.$$

Then $p(y)$ is 0 for $y = 0$, increases to a maximum 1 at $y = 1$, and then decreases to 0. It is $O(y)$ for small, and $O(y^{-1})$ for large y, so that the series

$$S = \sum_{\kappa=1}^{\infty} p(c^{-\kappa}), \qquad S' = \sum_{\kappa=1}^{\infty} p(c^{\kappa})$$

are convergent. Also

$$\sum_{r=0}^{R} |p_r| = \{4(1+1)\}^R = 8^R. \tag{7.13.5}$$

Suppose now that a_m is the a_n (or one of the a_n) whose modulus is largest. Then

$$\begin{aligned} \left|F\left(\frac{1}{\lambda_m}\right)\right| &= \left|\sum_{n=0}^{N} a_n P\left(\frac{\lambda_n}{\lambda_m}\right)\right| \\ &= \left|\sum_{n=0}^{N} a_n \left\{p\left(\frac{\lambda_n}{\lambda_m}\right)\right\}^R\right| \geqslant |a_m| - |a_m| \sum_{n \neq m} \left\{p\left(\frac{\lambda_n}{\lambda_m}\right)\right\}^R \\ &\geqslant |a_m|\left(1 - \sum_{\kappa=1}^{\infty} \{p(c^{-\kappa})\}^R - \sum_{\kappa=1}^{\infty} \{p(c^{\kappa})\}^R\right), \end{aligned} \tag{7.13.6}$$

since $\lambda_{n+1}/\lambda_n \geqslant c$ and $p(y)$ decreases as we move away from 1 in either direction. Also

$$\sum \{p(c^{-\kappa})\}^R \leqslant \{p(c^{-1})\}^{R-1}S, \qquad \sum \{p(c^{\kappa})\}^R \leqslant \{p(c)\}^{R-1}S',$$

each of which decreases as R increases and tends to 0 when $R \to \infty$. We can therefore choose an $R = R(c)$ for which $\sum \{p(c^{-\kappa})\}^R < \frac{1}{4}$, $\sum \{p(c^{\kappa})\}^R < \frac{1}{4}$,

$$|F(1/\lambda_m)| \geqslant |a_m|(1-\tfrac{1}{4}-\tfrac{1}{4}) = \tfrac{1}{2}|a_m|. \tag{7.13.7}$$

It now follows from (7.13.4), (7.13.5), and (7.13.7) that

$$|a_m| \leqslant 2.8^R H.$$

This proves Theorem 115, but we need its extension to infinite series.

THEOREM 116. *The result of Theorem* 115 *is true for an infinite series* $f(y) = \sum a_n e^{-\lambda_n y}$ *convergent for all positive* y.

We choose a particular n, say, $n = m$, and a positive ϵ. Then the series for $f(y+\epsilon)$ converges uniformly for $y > 0$, and we can choose $N = N(m, \epsilon) > m$ so that

$$\left|\sum_{n=1}^{N} a_n e^{-\lambda_n \epsilon} e^{-\lambda_n y}\right| < |f(y+\epsilon)| + \epsilon \leqslant H + \epsilon$$

for $y \geqslant 0$. Hence $|a_m e^{-\lambda_m \epsilon}| \leqslant C(H+\epsilon)$, and the result follows when $\epsilon \to 0$.

It is now easy to prove Theorem 114. Since $S(y) \to s$ when $y \to 0$, there is a $\delta = \delta(\epsilon)$ such that $|S(y)-S(y')| < \epsilon$ when y and y' both lie in $(0, 2\delta)$. Since $S(y)$ is continuous for $y > \delta$, and $S(y) \to 0$ when $y \to \infty$, there is an $\eta = \eta(\epsilon, \delta) = \eta(\epsilon)$ such that $0 < \eta < \delta$ and

$$|S(y)-S(y+\eta)| < \epsilon \tag{7.13.8}$$

for $y \geqslant \delta$; and this is true also for $0 < y < \delta$, since then y and $y+\eta$ both lie in $(0, 2\delta)$. Hence (7.13.8) is true for $y > 0$. Since

$$S(y)-S(y+\eta) = \sum a_n(1-e^{-\lambda_n \eta})e^{-\lambda_n y},$$

it follows from Theorem 116 that

$$|a_n|(1-e^{-\lambda_n \eta}) \leqslant C\epsilon$$

for all n, and so that $|a_n| \leqslant 2C\epsilon$ for large n. Hence $a_n = o(1)$. But then $a_n = o(\mu_n)$, because of (7.13.3), and the conclusion follows from Theorem 89.

NOTES ON CHAPTER VII

§ 7.1. A great deal has been written about Tauberian theorems during the last thirty years, and the literature is rather confusing, since almost every theorem carries a number of variants, analogues, and generalizations, and it is often difficult to trace a proof, or even an explicit statement, of the precise theorem which one may need. We confine ourselves here to theorems of 'power series type', i.e. theorems associated with the exponential kernel e^{-xy}, and to the simplest and most striking among them.

Our treatment of the subject in this chapter is based mainly on the work of Hardy and Littlewood and of Karamata. We return to it in Ch. XII, where we adopt the more general point of view of Wiener. There is a clear account of the fundamental theorems in Widder's ch. 5. The following list of papers may be useful:—

Ananda Rau [**1**], *JLMS*, 3 (1928), 200–5; [**2**], *PLMS* (2), 30 (1930), 367–72; [**3**], *RP*, 54 (1929), 455–61;

Bosanquet [**4**], *JLMS*, 19 (1944), 161–8;

Doetsch [**5**], *MA*, 82 (1921), 68–82;

Hardy and Littlewood [**6**], *PLMS* (2), 11 (1912), 411–78; [**7**], ibid. 13 (1913), 174–91; [**8**], ibid. 25 (1926), 219–36; [**9**], ibid. 30 (1930), 23–37; [**10**], *MM*, 43 (1914), 134–47;

Ingham [**11**], *OQJ*, 8 (1937), 1–7;

Karamata [**12**], *MZ*, 32 (1930), 319–20; [**13**], ibid. 33 (1931), 294–300; [**14**], *JM*, 164 (1931), 27–40;

Landau [**15**], *Monatshefte für Math.* 18 (1907), 8–28; [**16**], *RP*, 35 (1913), 265–76;

Littlewood [**17**], *PLMS* (2), 9 (1910), 434–48;

Rajagopal [**18**], *Math. Gazette*, 30 (1946), 272–6;

R. Schmidt [**19**], *MZ*, 22 (1925), 89–152;

Szász [**20**], *Münchener Sitzungsberichte* (1929), 325–40; [**21**], *TAMS*, 39 (1936), 117–30;

Tauber [**22**], *Monatshefte für Math.* 8 (1897), 273–7;

Titchmarsh [**23**], *PLMS* (2), 26 (1927), 185–200;

Vijayaraghavan [**24**], *JLMS*, 1 (1926), 113–20; [**25**], ibid. 2 (1927), 215–22.

The list is not complete, and does not include papers based on Wiener's ideas.

§ 7.2. Tauber [**22**]. The integral analogue, for the more general integral $\int \phi(yt)a(t)\,dt$, where $\phi'(t)$ is bounded, $\phi(0) = 1$, and $\int |\phi(t)|\,dt$ convergent, was proved by Hardy, *TCPS*, 21 (1910), 427–51 (432).

§ 7.3. Tauber [**22**]. The form of Theorem 88, with Stieltjes integrals, is that in which it is proved by Widder, 187, Theorem 3*b*.

§ 7.4. Theorem 89 was proved by Landau [**15**].

§ 7.5. Theorem 90 was proved, and Theorem 92 stated, by Littlewood [**17**]. The remaining theorems are due to Hardy and Littlewood [**7**]: all of them are proved in more general forms. There are generalizations for Dirichlet's series $\sum a_n e^{-\lambda_n s}$ in [**10**].

§ 7.6. Theorem 98 was proved by Szász [**20**]: it is the case $\gamma = 1$ of Widder's Theorem 4.3 (192). The proof here, based on Theorems 99 and 100, is substantially that of Karamata [**14**]. Theorem 96*a* was first proved explicitly (with a change of variable) by Doetsch [**5**]: see also Hardy and Littlewood [**9**] and Titchmarsh [**23**]. Doetsch also proves theorems equivalent to 91*a* and 94*a*.

§ 7.7. The first explicit proof of Theorem 101, in the form given here, seems to be that of Landau, *Ergebnisse*, 58. The theorem is stated and used by Hardy and Littlewood, [7] and [10]. The less general form in which $f''(y) = O(y^{-2})$ is included in Theorem 2 of [6] (420).

Theorem 102 is a slight generalization of Widder's Theorem 4.5 (195): he has $\beta(t) = t$.

Theorem 103 is proved by Hardy and Littlewood [10]. The example showing the necessity of the condition (7.7.12) is due to Ananda Rau [2].

When λ_n satisfies (7.7.12), Theorem 104 becomes the main theorem of Littlewood [17]. Littlewood says there that it is true without the restriction (7.7.12), but the first published proof of this is that of Ananda Rau [1]. We may complete the proof as follows.

We suppose $a_0 = 0$. Then, first, if

(1) $$a_n = O\{(\lambda_n - \lambda_{n-1})/\lambda_n\},$$

we have

(2) $$\sum_{m=1}^{n} \lambda_m a_m = O\left\{\sum_{m=1}^{n} (\lambda_m - \lambda_{m-1})\right\} = O(\lambda_n).$$

Secondly, by Theorem 88 [O], (2), together with $S(y) \to s$, implies

$$A(x) = \sum_{\lambda_n \leqslant x} a_n = O(1).$$

Next $$S(y) = \sum a_n e^{-\lambda_n y} = \int e^{-yt}\, dA(t) = y \int A(t) e^{-yt}\, dt,$$

since $A(0) = 0$, and so $$\int \{A(t)+H\} e^{-yt}\, dy \sim \frac{s+H}{y}$$

for any H. Choosing H so that $A(t)+H > 0$, and applying Theorem 96a, we find that

$$\int_0^t \{A(u)+H\}\, du \sim (s+H)t, \qquad \int_0^t A(u)\, du \sim st.$$

Finally, the conclusion follows from Theorem 67. This form of the proof is due to Bosanquet.

Szász [20, 21] proved that $\sum a_n$ converges to s if $S(y) \to s$ and a_n satisfies both (7.7.13) and (a) $\underline{\lim}\, a_n \geqslant 0$. This theorem includes Theorem 103, since (7.7.13) implies (a) when λ_n satisfies (7.7.12), and also the theorem referred to in the note on § 6.1.

Dr. Bosanquet has pointed out to me that (as was suggested to him by Mr. Ingham) (7.7.13) and $S(y) \to s$ imply

$$\sum a_n = s \ (\mathrm{R}, \lambda, \kappa)$$

for every positive κ; and Rajagopal [18] has proved this explicitly for $\kappa = 1$. Both Bosanquet and Rajagopal use a result of Szász [20], and Bosanquet also uses the theorem of Riesz for $(\mathrm{R}, \lambda, \kappa)$ summability which corresponds to Theorem 70.

Szász [20] and Ananda Rau [3] have proved that if $\sum a_n e^{-\lambda_n y} \sim y^{-\alpha}$, where $\alpha > 0$ and $a_n \geqslant 0$, then λ_n necessarily satisfies (7.7.12).

§ 7.10. Theorem 105 was proved by Szász [21]: it includes his theorem referred to under § 7.7. The method used in this section is that referred to at the end of Hardy and Littlewood [9].

§ 7.11. The proof is substantially that of Karamata [14].

§ 7.12. The technique of repeated differentiation was used first by Littlewood in [**17**].

§ 7.13. Theorem 114 was conjectured by Littlewood in [**17**], and proved by Hardy and Littlewood in [**8**]. The proof given here, which is much shorter, is due to Ingham [**11**]. Ingham proves a good deal more, in particular that, when $\lambda_{n+1}/\lambda_n \to \infty$, the limits of indetermination of s_n, when $n \to \infty$, are the same as those of $S(y)$ when $y \to 0$.

Bosanquet [**4**] has proved a theorem which includes both of Theorems 104 and 114, viz. that $S(y) \to s$ and

$$\lim_{\delta\to 0} \overline{\lim_{n\to\infty}} \max_{\lambda_n<\lambda_m<(1+\delta)\lambda_n} |a_{n+1}+...+a_m| = 0$$

imply $\sum a_n = s$. Szász [**21**] had proved the corresponding theorem for $(\mathrm{R}, \lambda, 1)$ summability.

VIII

THE METHODS OF EULER AND BOREL (1)

8.1. Introduction. In this and the next chapter we study more systematically a group of methods of which the most important are the E and B methods defined in §§ 1.3, 4.6, and 4.12–13. The definitions which we consider differ widely in form, and might seem at first sight unlikely to have much in common; but the relations between them turn out to be much closer than might have been expected. In particular the Tauberian theorems associated with them are essentially the same.

8.2. The (E, q) method. We begin with a generalization of the definition of §§ 1.3 and 4.6. Suppose that the series $\sum a_n x^{n+1}$ converges to $f(x)$ for small x, that $q > 0$, and that

(8.2.1) $$x = \frac{y}{1-qy}, \qquad y = \frac{x}{1+qx},$$

so that $y = (1+q)^{-1}$ when $x = 1$. Then, for small x and y,

(8.2.2) $$f(x) = \sum_0^\infty a_n\left(\frac{y}{1-qy}\right)^{n+1} = \sum_{n=0}^\infty a_n \sum_{m=n}^\infty \binom{m}{n} q^{m-n} y^{m+1}$$

$$= \sum_{m=0}^\infty y^{m+1} \sum_{n=0}^m \binom{m}{n} q^{m-n} a_n = \sum_0^\infty a_m^{(q)}\{(q+1)y\}^{m+1},$$

where

(8.2.3) $$a_m^{(q)} = \frac{1}{(q+1)^{m+1}} \sum_{n=0}^m \binom{m}{n} q^{m-n} a_n.$$

If

(8.2.4) $$\sum a_m^{(q)} = A,$$

then we say that $\sum a_n$ is summable (E, q) to sum A. For $q = 1$ the definition reduces to Euler's definition of §§ 1.3 and 4.6, and for $q = 0$ to that of ordinary convergence.

If $a_n = z^n$, then

$$a_m^{(q)} = \frac{(q+z)^m}{(q+1)^{m+1}}, \qquad \sum a_m^{(q)} = \frac{1}{q+1}\left(1-\frac{q+z}{q+1}\right)^{-1} = \frac{1}{1-z},$$

if and only if $|q+z| < q+1$. Thus $\sum z^n$ is summable (E, q) in the circle whose centre is $-q$ and whose radius is $q+1$. The circle increases with q, and tends to the half-plane $\Re z < 1$ when $q \to \infty$. We saw in §§ 4.12–13 that this is the region of B, or B′, summability of the series.

We may write (8.2.3) as

$$(8.2.5)\qquad (q+1)^{m+1}a_m^{(q)} = (q+E)^m a_0,$$

where E is defined as in § 4.6. Also

$$\frac{1}{q+1}+\frac{q+x}{(q+1)^2}+\dots+\frac{(q+x)^m}{(q+1)^{m+1}} = \frac{1}{(q+1)^{m+1}}\,\frac{(q+1)^{m+1}-(q+x)^{m+1}}{1-x}$$

$$= \frac{1}{(q+1)^{m+1}}\sum_{n=1}^{m+1}\binom{m+1}{n}q^{m+1-n}(1+x+x^2+\dots+x^{n-1}).$$

Hence, writing E for x, and observing that

$$(1+E+\dots+E^{n-1})a_0 = a_0+a_1+\dots+a_{n-1} = A_{n-1},$$

we obtain

$$(8.2.6)\qquad A_m^{(q)} = \sum_{n=0}^{m} a_n^{(q)} = \left\{\frac{1}{q+1}+\frac{q+E}{(q+1)^2}+\dots+\frac{(q+E)^m}{(q+1)^{m+1}}\right\}a_0$$

$$= \frac{1}{(q+1)^{m+1}}\left\{\binom{m+1}{1}q^m A_0+\binom{m+1}{2}q^{m-1}A_1+\dots+A_m\right\}.$$

There is a slight lack of symmetry in this formula which is inconvenient and will lead us to modify it in § 8.3.

We call $A^{(q)} = \sum a_n^{(q)}$ the q-th *Euler transform* of $A = \sum a_n$. The formal relation between the two series is defined by

$$\sum a_n x^{n+1} = \sum a_n^{(q)}\{(q+1)y\}^{n+1} = \sum a_n^{(q)} z^{n+1}, \qquad x = \frac{z}{1+q-qz}.$$

8.3. Simple properties of the (E, q) method. We must first prove

THEOREM 117. *The* (E, q) *method is regular.*

For, in the notation of § 3.2,

$$c_{m,n} = \frac{1}{(q+1)^{m+1}}\binom{m+1}{n+1}q^{m-n} > 0 \quad (n \leqslant m), \qquad c_{m,n} = 0 \quad (n > m),$$

$c_{m,n} \to 0$, and $\sum c_{m,n} = 1-(q+1)^{-m-1}q^{m+1} \to 1$ when $m \to \infty$.

Theorem 117 is the particular case $q' = 0$ of

THEOREM 118. *If a series is summable* (E, q'), *and* $q > q'$, *then it is summable* (E, q) *to the same sum.*

This plainly follows from Theorem 117 and

THEOREM 119. *The r-th Euler transform of the q-th Euler transform of a series is the* $(q+r+qr)$*-th Euler transform of the series.*

For $x = z/(1+q-qz)$ and $z = w/(1+r-rw)$ imply

$$x = \frac{w}{1+s-sw}, \qquad s = q+r+qr.$$

It follows from Theorem 118 that, as q increases, the (E, q) methods form a scale of increasing strength.†

THEOREM 120. *The* (E, q) *method has the properties* $(\alpha)-(\delta)$ *of Theorem* 40.

We need only consider (γ) and (δ). We have to show the equivalence of the two assertions

(8.3.1) $\sum a_n^{(q)} = A,$ (8.3.2) $\sum b_n^{(q)} = A - a_0,$

where $b_n = a_{n+1}$. We may suppose $a_0 = 0$, so that $B_n = A_{n+1}$. Then, after (8.2.6),

$$B_m^{(q)} = \frac{1}{(q+1)^{m+1}}\left\{\binom{m+1}{1}q^m A_1 + \binom{m+1}{2}q^{m-1}A_2 + \ldots + A_{m+1}\right\},$$

and so

(8.3.3)

$$B_m^{(q)} - A_m^{(q)} = \frac{1}{(q+1)^{m+1}}\left\{\binom{m+1}{1}q^m a_1 + \ldots + a_{m+1}\right\} = (q+1)a_{m+1}^{(q)}.$$

(i) If (8.3.1) is true, then $a_{m+1}^{(q)} \to 0$, and (8.3.2) follows from (8.3.3).

(ii) We may write (8.3.3) as

$$B_m^{(q)} = (q+1)A_{m+1}^{(q)} - qA_m^{(q)},$$

and it follows, since $A_0^{(q)} = 0$, that

$$(q+1)A_{m+1}^{(q)} = B_m^{(q)} + \frac{q}{q+1}B_{m-1}^{(q)} + \ldots + \left(\frac{q}{q+1}\right)^m B_0^{(q)}.$$

This is a transformation

$$A_{m+1}^{(q)} = \sum c_{m,n} B_n^{(q)}$$

with $c_{m,n} = q^{m-n}(q+1)^{-m+n-1} \quad (n \leqslant m),\ 0\,(n > m),$

and we can verify at once that the conditions of Theorem 2 are satisfied. Hence (8.3.2) and (8.3.3) imply $A_{m+1}^{(q)} \to A$, which is (8.3.1).

It follows from Theorem 120 that $A_n \to A$ (E, q) is equivalent to $A_{n+1} \to A$ (E, q), and so to

$$\frac{1}{(q+1)^{m+1}}\left\{q^{m+1}A_0 + \binom{m+1}{1}q^m A_1 + \ldots + A_{m+1}\right\} \to A.$$

Hence, changing $m+1$ into m, we may replace $A_m^{(q)} \to A$ by

$$\mathrm{A}_m^{(q)} = \frac{1}{(q+1)^m}\left\{q^m A_0 + \binom{m}{1}q^{m-1}A_1 + \ldots + A_m\right\} \to A;$$

and it is usually most convenient to define the 'Euler mean' of A_n in this way. We may say that $A_n \to A$ (E, q) if

(8.3.4) $$\mathrm{A}_m^{(q)} = \frac{1}{(q+1)^m}\sum_{n=0}^{m}\binom{m}{n}q^{m-n}A_n = \left(\frac{q+E}{q+1}\right)^m A_0 \to A.$$

† The example of the series $\sum z^n$ shows that no two (E, q) methods are equivalent.

If then we write s_n and t_n for A_n and $A_n^{(q)}$, we have

$$(8.3.5)\quad \Delta^m t_0 = \sum_{n=0}^{m} (-1)^n \binom{m}{n} t_n = \sum_{n=0}^{m} (-1)^n \binom{m}{n} \left(\frac{q+E}{q+1}\right)^n s_0$$

$$= \left(1-\frac{q+E}{q+1}\right)^m s_0 = \left(\frac{1-E}{q+1}\right)^m s_0 = \frac{1}{(q+1)^m}\Delta^m s_0,$$

an equation whose full significance will appear in Ch. XI.

THEOREM 121. *If $\sum a_n$ is summable* (E, q), *then* $a_n = o\{(2q+1)^n\}$.

It follows from (8.2.4) that $(q+1)a_m^{(q)} = o(1)$, and so, from (8.2.5), that

$$(q+E)^m a_0 = o\{(q+1)^m\}.$$

Also $a_n = E^n a_0 = (E+q-q)^n a_0$, and therefore

$$a_n = o\left\{(q+1)^n + \binom{n}{1} q(q+1)^{n-1} + \ldots + q^n\right\} = o\{(2q+1)^n\}.$$

The example of the series $\sum z^n$, which is summable (E, q) for

$$-2q-1 < z < 1,$$

shows that we cannot replace $2q+1$ by any smaller number.

8.4. The formal relations between Euler's and Borel's methods. We saw in § 8.2 that the region of (E, q) summability of $\sum z^n$ tends to its region of Borel summability when $q \to \infty$; and this suggests that Borel's method may be regarded as in some sense a limiting case of Euler's. We shall make this connexion more precise later (Theorem 128); but it may be worth while to show here how it harmonizes with the formal ideas of §§ 4.18 and 8.3.

If we write m/x for q in (8.3.4), we obtain

$$A_m^{(q)} = \left(1+\frac{m}{x}\right)^{-m} \left\{\left(\frac{m}{x}\right)^m A_0 + \binom{m}{1}\left(\frac{m}{x}\right)^{m-1} A_1 + \binom{m}{2}\left(\frac{m}{x}\right)^{m-2} A_2 + \ldots + A_m\right\}$$

$$= \left(1+\frac{x}{m}\right)^{-m} \left\{A_0 + xA_1 + \left(1-\frac{1}{m}\right)\frac{x^2}{2!}A_2 + \left(1-\frac{1}{m}\right)\left(1-\frac{2}{m}\right)\frac{x^3}{3!}A_3 + \ldots + \right.$$

$$\left. + \left(1-\frac{1}{m}\right)\left(1-\frac{2}{m}\right)\ldots\left(1-\frac{m-1}{m}\right)\frac{x^m}{m!}A_m\right\} = \phi(m, x),$$

say. Then

$$\lim_{m\to\infty} \lim_{x\to\infty} \phi(m,x) = \lim_{m\to\infty} A_m,$$

$$\lim_{x\to\infty} \lim_{m\to\infty} \phi(m,x) = \lim_{x\to\infty} e^{-x} \sum \frac{x^n}{n!} A_n.$$

The first way of proceeding to the limit leads to ordinary convergence, the second to Borel's exponential method of summation (§ 4.12). The various Euler methods correspond to the limit process $m = qx \to \infty$.

8.5. Borel's methods. Borel's exponential and integral methods were defined in §§ 4.12–13. If

$$e^{-x} \sum A_n \frac{x^n}{n!} \to A$$

we say that $A_n \to A$ (B), and if

$$\int_0^\infty e^{-x} \sum_0^\infty a_n \frac{x^n}{n!}\, dx = \lim_{X\to\infty} \int_0^X e^{-x} \sum_0^\infty a_n \frac{x^n}{n!}\, dx = A$$

we say that $A_n \to A$ (B′). The methods are of quite different types, the first being an 'integral function' definition in the sense of § 4.12, with $J(x) = e^x$, and the second a 'moment method' in the sense of § 4.13, with $\mu_n = n!$, $\chi(x) = 1-e^{-x}$; but the special properties of the exponential function make them all but equivalent. First, however, we observe

THEOREM 122. *The* B *and* B′ *methods are regular.*

This is a corollary of Theorem 33 (for B) and of Theorem 34 (for B′).

We now consider the relations between the two methods: we shall find that they are nearly, but not quite, equivalent. We write

$$a(x) = \sum a_n \frac{x^n}{n!}, \qquad A(x) = \sum A_n \frac{x^n}{n!}. \tag{8.5.1}$$

If one series is convergent for all x, then so is the other. Also

$$a'(x) = \sum a_{n+1} \frac{x^n}{n!}, \qquad A'(x) = \sum A_{n+1} \frac{x^n}{n!}, \tag{8.5.2}$$

$$\int_0^x e^{-t} a'(t)\, dt = e^{-x} a(x) - a_0 + \int_0^x e^{-t} a(t)\, dt, \tag{8.5.3}$$

$$e^{-x} A(x) - a_0 = \int_0^x \frac{d}{dt}\{e^{-t} A(t)\}\, dt = \int_0^x e^{-t}\{A'(t) - A(t)\}\, dt \tag{8.5.4}$$

$$= \int_0^x e^{-t} \sum_0^\infty (A_{n+1} - A_n) \frac{t^n}{n!}\, dt = \int_0^x e^{-t} \sum_0^\infty a_{n+1} \frac{t^n}{n!}\, dt = \int_0^x e^{-t} a'(t)\, dt;$$

and hence, comparing (8.5.3) and (8.5.4),

$$e^{-x} A(x) = e^{-x} a(x) + \int_0^x e^{-t} a(t)\, dt. \tag{8.5.5}$$

The last equation gives

THEOREM 123. *The* B *and* B′ *methods are equivalent if and only if* $e^{-x}a(x) \to 0$.

We can, however, go farther. It follows from (8.5.5) that, if

$$\int_0^x e^{-t}a(t)\,dt = \phi(x),$$

then $e^{-x}A(x) = \phi(x)+\phi'(x)$. If $\phi+\phi' \to A$ then, by Theorem 53, $\phi' \to 0$ and $\phi \to A$. Hence we deduce

THEOREM 124. *A series summable* (B) *is summable* (B′) *to the same sum.*

The converse is false. If

$$a_n = \sum \frac{(-1)^p(2p+2)^n}{(2p+1)!},$$

then

$$a(x) = \sum \frac{(-1)^p}{(2p+1)!} e^{(2p+2)x} = e^x \sin e^x,$$

$$\int_0^\infty e^{-x}a(x)\,dx = \int_0^\infty \sin e^x\,dx = \int_1^\infty \frac{\sin u}{u}\,du,$$

but $e^{-x}a(x)$ does not tend to 0, so that the series $\sum a_n$ is not summable (B). Thus we have

THEOREM 125. *There are series summable* (B′) *but not summable* (B).

Next, we observe

THEOREM 126. *The assertions*

$$a_0+a_1+a_2+\ldots = A \quad (\mathrm{B}), \qquad a_1+a_2+a_3+\ldots = A-a_0 \quad (\mathrm{B}')$$

are equivalent.

THEOREM 127. *The* B *and* B′ *methods possess the properties* (α), (β), *and* (γ) *of Theorem* 40, *but not the property* (δ).

Theorem 126 follows from (8.5.4), and Theorem 127 from Theorems 124, 125, and 126.

We conclude this section with the theorem to which we referred in § 8.4.

THEOREM 128. *If* $\sum a_n$ *is summable* (E, q), *then it is summable* (B) *or* (B′) *to the same sum.*

For
$$e^{qx}\sum A_n \frac{x^n}{n!} = \sum \frac{(qx)^n}{n!} \sum A_n \frac{x^n}{n!} = \sum c_n \frac{x^n}{n!},$$

where
$$\frac{c_n}{n!} = \frac{A_n}{n!} + \frac{qA_{n-1}}{(n-1)!\,1!} + \frac{q^2A_{n-2}}{(n-2)!\,2!} + \ldots + \frac{q^nA_0}{n!},$$

so that $c_n = (q+1)^n \mathrm{A}_n^{(q)}$, in the notation of (8.3.4). Hence

$$e^{-x}A(x) = e^{-x}\sum A_n\frac{x^n}{n!} = e^{-(q+1)x}\sum \mathrm{A}_n^{(q)}\frac{\{(q+1)x\}^n}{n!}.$$

If $\sum a_n$ is summable (E, q) to A, then $\mathrm{A}_n^{(q)} \to A$, and so $e^{-x}A(x) \to A$, by Theorem 122. Thus the series is summable (B), and *a fortiori* summable (B′).

8.6. Normal, absolute, and regular summability. If

$$a_p + a_{p+1} + a_{p+2} + \dots$$

is summable (B) for every p, in which case, after Theorem 124, it is also summable (B′) for every p, and conversely, by Theorem 126, then we shall say that $\sum a_n$ is *normally* summable. For this, it is necessary and sufficient that $\sum a_n$ should be summable and that

$$e^{-x}a^{(p)}(x) = e^{-x}\sum a_{n+p}\frac{x^n}{n!} \to 0.$$

If Borel's integral is absolutely convergent, we shall say that $\sum a_n$ is *absolutely* summable. If the series $a_p + a_{p+1} + \dots$ is absolutely summable for every p, i.e. if $\int e^{-x}|a^{(p)}(x)|\,dx < \infty$ for every p, then we shall say that $\sum a_n$ is *regularly* summable. Our language here differs from that of Borel, who defines absolute summability as we have defined regular summability. In any case the definitions will not be very important here.

The series $$\sum \frac{(1+i)^{n+1}}{n+1}$$

is normally but not absolutely summable. Its sum is

$$\int\limits_0^\infty e^{-t}a(t)\,dt = \int\limits_0^\infty \frac{\cos t - e^{-t}}{t}\,dt + i\int\limits_0^\infty \frac{\sin t}{t}\,dt = \tfrac{1}{2}\pi i.$$

8.7. Abelian theorems for Borel summability. Our next theorems are 'Abelian': they belong to the class typified by Abel's theorem on the continuity of power series. Here, and throughout the rest of the chapter, we work primarily in terms of summability (B′): the transition to summability (B), when desirable, is easily effected by means of Theorem 126.

THEOREM 129. *If a power series $\sum a_n z^n$ is summable* (B′) *at a point P, then it is summable at every point of the stretch OP from the origin to P. If Q is a point on OP between O and P, then the series is uniformly summable on QP.*

It is not assumed that the series has a circle of convergence. We may suppose (making a trivial transformation) that P is the point $z = 1$. Then the integral

$$J(z) = \int e^{-t} a(zt)\,dt \tag{8.7.1}$$

is convergent for $z = 1$, and we have to prove it convergent for $0 < z \leqslant 1$ and uniformly convergent for $0 < \delta \leqslant z \leqslant 1$. Now

$$J(z) = \frac{1}{z} \int e^{-t/z} a(t)\,dt = \frac{K(z)}{z}, \tag{8.7.2}$$

say, when $0 < z \leqslant 1$; and

$$K(z) = \int e^{-st} e^{-t} a(t)\,dt = k(s), \qquad s = \frac{1}{z} - 1. \tag{8.7.3}$$

This integral is uniformly convergent for $s \geqslant 0$, i.e. for $0 < z \leqslant 1$, and therefore $J(z)$ converges as stated in the theorem.

Theorem 129 does not state the full truth; actually, $J(z)$ converges uniformly for $0 \leqslant z \leqslant 1$. The argument above fails to prove this because of the factor z^{-1} in (8.7.2).

THEOREM 130. *If $\sum a_n z^n$ is summable* (B′) *at P, then it is uniformly summable on OP.*

We may again suppose that P is $z = 1$; and it is also convenient to suppose, as plainly we may, that a_n is real. We have to prove that

$$|I| = |I(z, H, H')| = \left| \int_H^{H'} e^{-t} a(zt)\,dt \right| < \epsilon \tag{8.7.4}$$

for $H' > H \geqslant H_0(\epsilon)$ and $0 \leqslant z \leqslant 1$. Since Theorem 129 proves uniform convergence over $(\frac{1}{2}, 1)$, we may suppose $0 < z \leqslant \frac{1}{2}$. There are three cases to consider, according as (*a*) $H'z \leqslant 1$, (*b*) $Hz < 1 < H'z$, or (*c*) $Hz \geqslant 1$. We state the argument for case (*b*), the arguments in the other cases being simpler variants. We may suppose $H > 2$.

We write

$$M = \operatorname*{Max}_{0 \leqslant t \leqslant 1} |a(t)|, \qquad N = \operatorname*{Max}_{T > 1} \left| \int_1^T e^{-t} a(t)\,dt \right|.$$

Then

$$I = \int_H^{1/z} e^{-t} a(zt)\,dt + \int_{1/z}^{H'} e^{-t} a(zt)\,dt = I_1 + I_2, \tag{8.7.5}$$

$$|I_1| \leqslant M \int_H^{\infty} e^{-t}\,dt = Me^{-H}, \tag{8.7.6}$$

$$I_2 = \frac{1}{z} \int_1^{H'z} e^{-t/z} a(t)\,dt = \frac{1}{z} \int_1^{H'z} e^{-st} e^{-t} a(t)\,dt = \frac{e^{-s}}{z} \int_1^{T} e^{-t} a(t)\,dt,$$

where $s = z^{-1}-1$ and $1 < T < H'z$. Hence

$$(8.7.7)\qquad |I_2| \leqslant \frac{N}{z}\exp\left(1-\frac{1}{z}\right) \leqslant \frac{N}{z}e^{-1/2z} \leqslant NHe^{-\frac{1}{2}H},$$

since $0 < z \leqslant \frac{1}{2}$, $2 \leqslant H \leqslant 1/z$, and $ue^{-\frac{1}{2}u}$ decreases for $u > 2$. From (8.7.5)–(8.7.7) it follows that

$$|I| \leqslant Me^{-H}+NHe^{-\frac{1}{2}H} < \epsilon$$

for $H \geqslant H_0(\epsilon)$.

As an application of Theorem 130 we prove

THEOREM 131. *If $\sum a_n$ is summable* (B′), *and*

$$(8.7.8)\qquad c_n = \int\limits_0^1 x^n\,d\chi(x),$$

where $\chi(x)$ is bounded and increases with x, then $\sum c_n a_n$ is summable (B′).

For, if $b_n = c_n a_n$, then

$$b(t) = \sum b_n\frac{t^n}{n!} = \sum a_n\frac{t^n}{n!}\int\limits_0^1 x^n\,d\chi = \int\limits_0^1 a(tx)\,d\chi,$$

$$\int\limits_0^\infty e^{-t}b(t)\,dt = \int\limits_0^1 d\chi\int\limits_0^\infty e^{-t}a(tx)\,dt,$$

because the inner integral on the right is uniformly convergent for $0 \leqslant x \leqslant 1$.

THEOREM 132. *If $\sum a_n z^n$ is summable* (B′) *at P, then its sum on OP is an analytic function of z regular inside the circle C described on OP as diameter.*

We may again suppose that P is $z = 1$. The series is summable on OP, and its sum is given, for $0 < z < 1$, by (8.7.2). It is sufficient to prove that $K(z)$ converges uniformly in the region D bounded by any two circular arcs from O to P making acute angles η with OP. We write

$$z = re^{i\theta},\qquad s = z^{-1}-1 = \rho e^{i\phi}$$

and use the formula (8.7.3). Since $k(s)$ is convergent for $s = 0$, it is uniformly convergent in the angle $|\phi| \leqslant \eta$. The arms of this angle correspond to the circular arcs which bound D, and its interior to the interior of D. Hence $K(z)$ is uniformly convergent in D.

It will be observed that the transformation from (8.7.1) to (8.7.2) presupposes the reality of z. Thus, although we have proved that $J(z)$ is regular inside C, we have not proved the series summable except on OP; and we shall see later (§ 8.9) that it is not necessarily summable at any other point of C.

8.8. Analytic continuation of a function regular at the origin: the polygon of summability. If the series $\sum a_n z^n$ has a circle of convergence, it defines a function regular at the origin, and the integrals $J(z)$ and $K(z)$ of § 8.7 may be used to find representations of this

function valid outside the circle. We can define the region of convergence of $J(z)$ in terms of the singularities of the function.

The function $$f(z) = (c-z)^{-1} = \sum c^{-n-1}z^n$$

is regular except at $z = c$, or P, and its circle of convergence is $|z| = |c|$. In this case

$$J(z) = c^{-1}\int e^{-t(1-z/c)}\,dt$$

is convergent if and only if $\mathfrak{R}(z/c) < 1$, i.e. if z and the origin lie on the same side of the line L_P through P perpendicular to OP. If

$$f(z) = \sum_{m=1}^{n} \frac{\gamma_m}{c_m - z}, \tag{8.8.1}$$

and $z = c_m$ is P_m, then z must lie on the same side as the origin of all the lines L_{P_m}. The region thus defined is the inside of a convex polygon, which may be closed or open and may reduce to an angle, strip, or half-plane. The series is summable 'inside' this polygon. Cauchy's integral formula, which is a generalization of (8.8.1), suggests that there may be a corresponding result for an arbitrary analytic function regular at the origin.

We suppose then that $f(z) = \sum a_n z^n$ is regular at O, that P is a singular point of $f(z)$, and S the set of all points P. We define Π or $\Pi(f)$ as the set of all points Q such that Q and O lie on the same side of every L_P, Γ as the set of frontier points of Π, and Π^* as the part of the plane complementary to $\Pi+\Gamma$. We call Γ the *Borel polygon* of f, Π its interior, and Π^* its exterior; and we shall prove that Π is the region of summability of $\sum a_n z^n$ in the sense that the series is summable at all points of Π and is not summable at any point of Π^*.

If $f(z) = (1-z^2)^{-1}$, then Γ is formed by the two lines $x = \pm 1$ and Π is the strip between them. If the circle of convergence is a barrier of singularities, Γ coincides with it. If $x = d > 0$ is a barrier of singularities, and $f(z)$ is regular to the left of this line, then Γ is the parabola which touches the line at d and has O as focus.

It follows at once from Theorem 132 that the series is not summable at any point Q of Π^*. For, if Q belongs to Π^*, there is a line L_P passing between O and Q, and the corresponding P lies inside the circle C. It remains to prove that the series is summable at points of Π.

Suppose that $f(u)$ is regular in and on a closed curve K surrounding the origin in the u-plane, and that

$$\mathfrak{R}(z/u) \leqslant 1-\delta < 1 \tag{8.8.2}$$

for all points u on K (in which case z necessarily lies inside K). Then

$$(8.8.3) \qquad f(z) = \frac{1}{2\pi i}\int_K \frac{f(u)}{u-z}\,du = \frac{1}{2\pi i}\int_K \frac{f(u)}{u}\,du \int e^{-t+tz/u}\,dt.$$

The repeated integral is majorized by

$$\frac{1}{2\pi}\int_K \frac{|f(u)||du|}{|u|}\int e^{-\delta t}\,dt.$$

We may therefore invert the integrations; and we obtain

$$(8.8.4) \qquad f(z) = \int e^{-t}\,dt\,\frac{1}{2\pi i}\int_K \frac{f(u)}{u}e^{tz/u}\,du = \int e^{-t}I(t,z)\,dt,$$

say. Since $f(u)$ is regular inside K, and $e^{tz/u}$ regular except at the origin, we can calculate $I(t,z)$ by contracting K into a curve K' inside the circle of convergence of $f(u)$. The power series for $f(u)$ and $e^{tz/u}$ are uniformly convergent on K', and so

$$I(t,z) = \frac{1}{2\pi i}\int_{K'} \sum a_n u^n \sum \frac{1}{n!}\left(\frac{tz}{u}\right)^n \frac{du}{u} = \sum a_n \frac{(tz)^n}{n!} = a(tz).$$

Hence $$f(z) = \int e^{-t}a(tz)\,dt,$$

i.e. $\sum a_n z^n$ is summable to $f(z)$.

It remains to show that, if z is in Π, we can draw K so as to satisfy our conditions. If Q is a point of Π, then $f(z)$ is regular inside the circle C described on OQ as diameter; for, if there were a singular point P inside C, the corresponding L_P would pass between O and Q. Further, since there are points Q' of Π on OQ beyond Q, and $f(z)$ is regular at O, it is regular on a slightly larger concentric circle C' intersecting OQ in O' and Q'. If z is at Q and u at a point A of C', then $\Re(z/u) < 1$ if Q and O lie on the same side of the line through A perpendicular to OA. The envelope of these lines, when A runs round C', is an ellipse whose foci are O and Q and whose major axis is $O'Q'$: C' is the 'auxiliary circle' of the ellipse, which is flat when O' is near to O, but always includes OQ in its interior. Also $\Re(z/u) < 1$ for all u on C', and therefore, since $\Re(z/u)$ is continuous, (8.8.2) is satisfied, with an appropriate δ, for all u on C'. It follows that, when z is at Q, we can take C' as the curve K of our argument, and therefore that the series is summable at Q. Thus it is summable at any point of Π.

Our argument actually proves rather more. The repeated integral (8.8.3) is absolutely convergent; and therefore (8.8.4) is absolutely con-

vergent, so that the series is absolutely summable at Q. Also the function $f_p(z) = a_p z^p + a_{p+1} z^{p+1} + \dots$ satisfies the same conditions of regularity as $f(z)$, so that all the series $a_p z^p + \dots$ are absolutely summable. Hence $\sum a_n z^n$ is regularly summable at Q.

It is also plain that the whole argument works uniformly for all z in any closed region interior to Π, so that the series is uniformly summable in any such region.

Summing up, we have proved

THEOREM 133. *The power series representing a function regular at the origin is summable* (B′) *inside the Borel polygon of the function, regularly, and uniformly throughout any closed region interior to the polygon; and is not summable at any point outside the polygon.*

In particular we have

THEOREM 134. *A power series is summable* (B′) *at any regular point on its circle of convergence, and uniformly summable in some neighbourhood of any such point.*

We may plainly substitute B for B′ in these theorems.

8.9. Series representing functions with a singular point at the origin. The analysis of § 8.8 rests throughout on the assumption that the series $\sum a_n z^n$ converges for small z. When this is not true, the series may still be summable for certain z, and give a complete or partial representation of an analytic function; but the region or regions of summability may have very diverse characters, and the sums in different regions may represent different functions. In all cases, however, after Theorem 130, a region of summability which includes a point P must include all of the line OP.

The two examples which follow are interesting.

(1) If the series is

$$1+0-\frac{2!}{1!}z^2+0+\frac{4!}{2!}z^4+0-\dots,$$

then $a(zt) = e^{-z^2t^2}$, and the sum is

$$J(z) = \int e^{-t-z^2t^2}\,dt. \tag{8.9.1}$$

If $z = re^{i\theta}$, then the integral converges in the quadrant $-\frac{1}{4}\pi < \theta < \frac{1}{4}\pi$ and its image with respect to the origin. If $z = x = 1/\xi > 0$, then

$$J(z) = \xi\int_0^\infty e^{-\xi u-u^2}\,du = \xi e^{\frac{1}{4}\xi^2}\int_{\frac{1}{2}\xi}^\infty e^{-v^2}\,dv = \xi e^{\frac{1}{4}\xi^2}\left(\tfrac{1}{2}\sqrt{\pi}-\int_0^{\frac{1}{2}\xi} e^{-v^2}\,dv\right) = F(\xi),$$

say, an integral function of ξ. Thus $J(z) = F(1/z)$ for $|\arg z| < \frac{1}{4}\pi$. Also $J(z)$ is even, so that $J(z) = F(-1/z)$ in the opposite quadrant; and these two functions differ by $\pi^{\frac{1}{2}}z^{-1}e^{1/4z^2}$. Thus the series represents different analytic functions in its two regions of summability.

(2) Suppose that
$$a_n = \sum \frac{(-1)^p}{p!} p^n c^p,$$
where $c > 0$. Then
$$a(zt) = \sum \frac{(zt)^n}{n!} \sum \frac{(-1)^p}{p!} p^n c^p$$
$$= \sum \frac{(-1)^p c^p}{p!} \sum \frac{(pzt)^n}{n!} = \sum \frac{(-1)^p c^p}{p!} e^{pzt} = e^{-ce^{zt}},$$
and
$$J(z) = \int e^{-t-ce^{zt}}\, dt = \int e^{-t-ce^{xt}(\cos yt + i \sin yt)}\, dt.$$
If $x \leqslant 0$ then $J(z)$ converges for all y, but if $x > 0$ it converges for $y = 0$ only. Thus the series is summable (1) in the half-plane $x \leqslant 0$, and (2) on the positive real axis.

Let us first suppose z real. Then
$$J(z) = \int\limits_0^\infty e^{-t-ce^{zt}}\, dt = \int\limits_1^\infty e^{-cu^z} \frac{du}{u^2}.$$
Putting $u^z = v$ and $Z = -1/z$, we find that $J(z)$ is $P(Z)$ or $Q(Z)$, where
$$P(Z) = -Z \int\limits_1^\infty e^{-cv} v^{Z-1}\, dv, \qquad Q(Z) = Z \int\limits_0^1 e^{-cv} v^{Z-1}\, dv,$$
according as $z > 0$ or $z < 0$. Here $P(Z)$ is an integral function of Z; and
$$Q(Z) = \Gamma(Z+1)c^{-Z} + P(Z)$$
if $\Re Z > 0$, so that $Q(Z)$ defines a function analytic and meromorphic all over the plane. The two functions represented by the series differ by $\Gamma(Z+1)c^{-Z}$.

This example is particularly interesting as an illustration of Theorems 130 and 132. If P is a point on the positive real axis, then the series is uniformly summable on OP, and is regular inside the circle C of Theorem 132, but it is not summable at any point in C except points on the axis. In this sense Theorem 130 states the most that is true.

8.10. Analytic continuation by other methods. The principles used in § 8.8 may be applied to other methods of summation. The most interesting for this purpose are those which, like Lindelöf's and Mittag-Leffler's methods of § 4.11, sum $\sum z^n$ in its Mittag-Leffler star. We consider, generally, a method P of summation in which $\sum a_n$ is defined as

(8.10.1)
$$\lim_{\delta \to 0} \sum A_n(\delta) a_n,$$
where $A_n(\delta) \to 1$, when $\delta \to 0$, for every n. Thus
$$A_0(\delta) = 1, \qquad A_n(\delta) = e^{-\delta n \log n} \quad (n > 0),$$
for Lindelöf's, and $A_n(\delta) = \{\Gamma(1+\delta n)\}^{-1}$ for Mittag-Leffler's method.

THEOREM 135. *Suppose that* (i) $\sum A_n(\delta) z^n$ *is an integral function of* z, *for every positive* δ; (ii) *that*
$$\phi_\delta(z) = \sum A_n(\delta) z^n \to \frac{1}{1-z}$$

when $\delta \to 0$, uniformly in any closed and bounded region containing no point of the line $(1,\infty)$; (iii) *that $f(z)$ is the principal branch of an analytic function regular at O and represented by $\sum a_n z^n$ for small z. Then*

$$\sum A_n(\delta) a_n z^n \to f(z)$$

uniformly in any closed and bounded region Δ interior to the Mittag-Leffler star of $f(z)$.

We may suppose that Δ is star-shaped, i.e. that if it includes P then it includes every point of OP. We can expand Δ about the origin, in a ratio $\rho > 1$, into a region Δ' still lying in the star of $f(z)$. If K is the boundary of Δ', and z is inside Δ, then z/u is not on $(1,\infty)$ for any u of K, and

$$\phi_\delta\left(\frac{z}{u}\right) \to \frac{u}{u-z}$$

uniformly on K. Hence

$$f(z) = \frac{1}{2\pi i}\int_K \frac{f(u)}{u-z}\,du = \frac{1}{2\pi i}\int_K \left\{\lim_{\delta\to 0}\phi_\delta\left(\frac{z}{u}\right)\right\} f(u)\frac{du}{u}$$

$$= \lim_{\delta\to 0}\frac{1}{2\pi i}\int_K \phi_\delta\left(\frac{z}{u}\right) f(u)\frac{du}{u} = \lim_{\delta\to 0}\frac{1}{2\pi i}\int_K \sum A_n(\delta)\left(\frac{z}{u}\right)^n f(u)\frac{du}{u}.$$

Contracting K into a contour inside the circle of convergence of $f(u)$, as in § 8.8, substituting the power series for $f(u)$, and integrating term by term, we obtain

$$f(z) = \lim_{\delta\to 0}\sum A_n(\delta) a_n z^n.$$

It is plain that the argument works uniformly for z of Δ. Methods such as these give better results in this problem than Borel's, but Borel's method is much more manageable, owing to the simple formal properties of the exponential function and series.

8.11. The summability of certain asymptotic series. It has been proved by Borel and Carleman that there are analytic functions corresponding to arbitrary asymptotic series (§ 2.5). More precisely, given any sequence (a_n) and any positive α, there are functions $f(z) = f(re^{i\theta})$ such that

$$f(z) \sim a_0 + a_1 z^{-1} + a_2 z^{-2} + \dots,$$

when $r \to \infty$, uniformly in the angle $|\theta| < \alpha\pi$. If $B > 0$ and $2k\alpha < 1$, then $r^n e^{-Bz^k} \to 0$ for every n, uniformly in the angle. Thus all the functions $f(z) + Ae^{-Bz^k}$ have the same asymptotic expansion in the angle.

The situation is changed if we adopt more precise hypotheses concerning the error-term of the expansion. It may then be possible to prove that there is at most one $f(z)$ which satisfies the conditions, and

that $f(z)$ is in some sense, for example Borel's, the sum of the series. We show this by proving a theorem of Watson.

We denote the region

$$r \geqslant k > 0, \qquad -\tfrac{1}{2}\pi-\mu \leqslant \theta \leqslant \tfrac{1}{2}\pi+\lambda,$$

where $0 < \lambda < \frac{1}{2}\pi$, $0 < \mu < \frac{1}{2}\pi$, by $D(\lambda, \mu, k)$, and its boundary by $C(\lambda, \mu, k)$; the latter is formed by part of a circle whose centre is the origin and parts of two radii from the origin to infinity in the negative half-plane. Usually λ and μ will be equal.

THEOREM 136. *Suppose that $f(z)$ is regular in $D(\lambda, \lambda, k)$, for a given λ and k, that $\sigma > 0$, and that*

$$f(z) = a_0+a_1 z^{-1}+\ldots+a_n z^{-n}+R_n(z), \tag{8.11.1}$$

where

$$a_n = O(n!\,\sigma^n), \qquad R_n = O\{(n+1)!\,\sigma^{n+1}r^{-n-1}\}, \tag{8.11.2}$$

uniformly for all n and for z in $D(\lambda, \lambda, k)$. Then (i) *the series*

$$\sum a_n \frac{t^n}{n!} = a(t), \tag{8.11.3}$$

where $t = \rho e^{i\phi}$, is convergent for $\rho < 1/\sigma$; (ii) *the function $a(t)$ is regular in any angle $|\phi| \leqslant \delta < \lambda$;*

(iii)
$$a(t) = \frac{1}{2\pi i}\int_L f\left(\frac{u}{t}\right)\frac{e^u}{u}\,du,$$

where L is a contour $C(\nu, \nu, l)$ with $0 < \nu < \lambda$ and $l > k/\sigma$, described from below; and

(iv)
$$f(z) = \int_0^\infty e^{-w} a\left(\frac{w}{z}\right) dw$$

for $r > k$, $|\theta| \leqslant \delta$.

The last clause of the theorem asserts that $\sum a_n z^{-n}$ is summable, to $f(z)$, by an extension of Borel's method which is often useful. If (*a*) the series (8.11.3) is convergent for small t, (*b*) the function $a(t)$ defined by the series is regular on $(0, \infty)$, and (*c*) $\int e^{-t}a(t)\,dt = s$, then we shall say that $\sum a_n$ is summable (B*) to s. Thus here $\sum a_n z^{-n}$ is summable (B*) to $f(z)$.

As another example, if $a_n = (-1)^n n!\,z^n$, and z is not real and negative, then $a(t) = (1+zt)^{-1}$ and

$$1-1!\,z+2!\,z^2-\ldots = \int \frac{e^{-t}}{1+zt}\,dt \quad (\mathrm{B}^*).$$

This is the sum found heuristically in § 2.4.

If the inequalities (8.11.2) are true for every positive σ, then $a(t)$ is an integral function, and B* reduces to B′.

Passing to the proof of Theorem 136, (i) is obvious from (8.11.2). Next, suppose that $t = \rho e^{i\phi}$, $\rho < 1/\sigma$. Then, replacing $f(u/t)$ by its asymptotic expansion, we have formally

$$\text{(8.11.4)}\qquad \frac{1}{2\pi i}\int\limits_L f\left(\frac{u}{t}\right)\frac{e^u}{u}\,du = \sum a_m t^m \frac{1}{2\pi i}\int\limits_L \frac{e^u}{u^{m+1}}\,du = \sum a_m \frac{t^m}{m!},$$

and we have to justify the term-by-term integration. If u describes L then u/t describes $C(\nu-\phi, \nu+\phi, l/\rho)$, which lies inside $D(\lambda, \lambda, k)$ if

$$\text{(8.11.5)}\qquad l > k/\sigma > k\rho, \qquad |\phi| < \lambda-\nu.$$

Then $f(u/t)$ is bounded on L, and the integrals in (8.11.4) are convergent.

We now write the integral (8.11.4) in the form

$$\text{(8.11.6)}\quad \sum_{m=0}^{n} a_m t^m \frac{1}{2\pi i}\int\limits_L \frac{e^u}{u^{m+1}}\,du + \frac{1}{2\pi i}\int\limits_L R_n\left(\frac{u}{t}\right)\frac{e^u}{u}\,du = \sum_{m=0}^{n} a_m \frac{t^m}{m!} + P_n,$$

say, and find an upper bound for $|P_n|$ for large n. We may suppose $n > l > k/\sigma$. Since $R_n(z)$ is regular inside $D(\lambda, \lambda, k)$, we may increase the radius of the circular part of L from l to n. Then the contribution of the circular part is

$$O\{(n+1)!\,(\sigma\rho)^{n+1}e^n n^{-n-1}\} = O\{n^{\frac{1}{2}}(\sigma\rho)^{n+1}\},$$

and that of the rectilinear parts is

$$O\left\{(n+1)!\,(\sigma\rho)^{n+1}\int\limits_n^\infty \frac{e^{-r\sin\nu}}{r^{n+2}}\,dr\right\} = O\{(n+1)!\,(\sigma\rho)^{n+1}n^{-n-2}\int\limits_0^\infty e^{-r\sin\nu}\,dr$$

$$= O\left\{\frac{(n+1)!\,(\sigma\rho)^{n+1}}{n^{n+2}\sin\nu}\right\} = O\{n^{-\frac{1}{2}}e^{-n}(\sigma\rho)^{n+1}\}.$$

Thus $P_n \to 0$, and (8.11.6) gives (8.11.4), subject to the conditions (8.11.5).

Suppose now that t varies in any region T defined by $|\phi| \leqslant \delta < \lambda$, $0 < \rho_1 \leqslant \rho \leqslant \rho_2$. Then we can choose ν and l so that the conditions (8.11.5) are satisfied for all t of T, and the integral (8.11.4) is then uniformly convergent in T, so that it defines a function $a(t)$ regular in T. We have thus proved (ii) and (iii).

If w is positive, $z = re^{i\theta}$, $r > k$, $|\theta| \leqslant \delta$, and $t = w/z = \rho e^{i\phi}$, then $|\phi| \leqslant \delta$, so that

$$a\left(\frac{w}{z}\right) = \frac{1}{2\pi i}\int\limits_L f\left(\frac{uz}{w}\right)\frac{e^u}{u}\,du.$$

When u describes L, $v = uz/w$ describes $C(\nu+\theta, \nu-\theta, l/\rho)$ or L', and

$$a\left(\frac{w}{z}\right) = \frac{1}{2\pi i}\int_{L'} f(v)\frac{e^{wv/z}}{v}\,dv.$$

We choose r_1 so that $k < r_1 < r$. Since $l/\rho > k/(\sigma\rho) > k$, we may replace L' by

$$L_1 = C(\nu+\theta, \nu-\theta, r_1).$$

We have then

$$\int e^{-w}a\left(\frac{w}{z}\right)dw = \int e^{-w}\,dw\frac{1}{2\pi i}\int_{L_1} f(v)\frac{e^{wv/z}}{v}\,dv$$

$$= \frac{1}{2\pi i}\int_{L_1}\frac{f(v)}{v}\,dv\int e^{-w(1-v/z)}\,dw = \frac{1}{2\pi i}\int_{L_1} f(v)\frac{z\,dv}{v(z-v)} = f(z),$$

provided that we may invert the integrations; and this is so if the double integral is absolutely convergent. We consider the circular and rectilinear parts of L_1 separately. On the circular part

$$\Re(1-v/z) \geqslant 1-r_1/r > 0,$$

so that its contribution is majorized by a multiple of $\int |v|^{-1}\,|dv| < 2\pi$. On the upper rectilinear part

$$\operatorname{am}\frac{wv}{z} = \operatorname{am} u = \tfrac{1}{2}\pi+\nu, \qquad \Re\left(\frac{wv}{z}\right) = -w\left|\frac{v}{z}\right|\sin\nu,$$

so that its contribution is majorized by a multiple of

$$\int e^{-w}\,dw\int e^{-w|v/z|\sin\nu}\frac{|dv|}{|v|} = \int\frac{1}{|v|}\frac{|z|\,|dv|}{|z|+|v|\sin\nu} < \infty;$$

and the lower rectilinear part may be dealt with similarly. Thus the double integral is absolutely convergent, and this completes the proof of the theorem.

Theorem 136 shows incidentally that at most one $f(z)$ can satisfy (8.11.1) and (8.11.2). But if we are concerned only with the uniqueness of $f(z)$, then we can prove more, and without reference to the theory of summability. We need only suppose that $f(z)$ satisfies (8.11.1) and (8.11.2) in the angle $|\theta| \leqslant \frac{1}{2}\pi$, instead of the larger region of Theorem 136. If $f_1(z)$ and $f_2(z)$ both satisfy (8.11.1) and (8.11.2) for $|\theta| \leqslant \frac{1}{2}\pi$, and $g(z) = f_1(z)-f_2(z)$, then

$$|g(z)| = O\{(n+1)!\,\sigma^{n+1}r^{-n-1}\}$$

uniformly in n and θ. We take $n = [r/\sigma]$, and a simple application of Stirling's theorem shows that $|g(z)| = O\{e^{-(1-\delta)r/\sigma}\}$ for every positive δ and $|\theta| \leqslant \frac{1}{2}\pi$. It follows, from familiar theorems of the Phragmén-Lindelöf type, that $g(z) = 0$.

Carleman has gone much farther, and found a necessary and sufficient condition that

$$|g(z)| \leqslant \alpha_n^n r^{-n} \quad (r \geqslant r_0 > 0, \quad |\theta| \leqslant \tfrac{1}{2}\pi)$$

should imply $g(z) = 0$. For suitably regular α_n, the condition is the divergence of $\sum \alpha_n^{-1}$. In our case this is effectively the harmonic series $\sum n^{-1}$.

It will be observed that we prove the summability of the series in an angle smaller (by π) than that in which it is supposed to represent $f(z)$ asymptotically. The example of § 8.9(1), with $1/z$ for z, shows that this is a real limitation corresponding to the facts of the case. There $a(t) = e^{-t^2}$, and the integral

$$f(z) = \int e^{-t-t^2/z^2}\,dt$$

converges if $|\theta| \leqslant \frac{1}{4}\pi$ but not if $\frac{1}{4}\pi < |\theta| < \frac{3}{4}\pi$. If z is positive,

$$f(z) = z\int e^{-uz-u^2}\,du = \sum \frac{(-1)^p z^{p+1}}{p!}\int e^{-u^2}u^p\,du$$

$$= \tfrac{1}{2}z\sqrt{\pi}\sum \frac{(-\frac{1}{2}z)^p}{\Gamma(1+\frac{1}{2}p)} = \tfrac{1}{2}z\sqrt{\pi}\,E_{\frac{1}{2}}(-\tfrac{1}{2}z),$$

where $E_{\frac{1}{2}}(z)$ is Mittag-Leffler's function. It is known that $f(z)$ has the asymptotic expansion

$$1-\frac{2!}{1!}\frac{1}{z^2}+\frac{4!}{2!}\frac{1}{z^4}-\dots$$

for $|\theta| < \frac{3}{4}\pi$, so that the series is asymptotic in an angle greater by π than that in which it is summable. We could enlarge the angle of validity of the integral representation of $f(z)$ by taking it along a line making an angle with the real axis.

NOTES ON CHAPTER VIII

§§ 8.2–4. The first systematic account of the (E, q) methods was given by Knopp, *MZ*, 15 (1922), 226–53, and 18 (1923), 125–56; and much of the argument of these sections is modelled on his. In particular Theorems 117–21 are Knopp's.

The (E, 1) method, and those derived from it by iteration, had been used frequently before, especially for purposes of numerical computation. Examples will be found in Bromwich, 62–6 and 196–8.

There are some passages in the first edition of Bromwich (302–10) which may seem at first to contradict some of the assertions here and in § 8.5. The explanation is that Bromwich, when he applies 'Euler's method' to power series, does not use the right definition. According to our definition $\sum a_n z^n$ is defined by

$$\tfrac{1}{2}a_0+\tfrac{1}{4}(a_0+a_1 z)+\tfrac{1}{8}(a_0+2a_1 z+a_2 z^2)+\dots.$$

Bromwich, in effect, uses the identity

$$a_0+a_1 z+a_2 z^2+\dots = a_0\frac{1}{1+z}+(a_0+a_1)\frac{z}{(1+z)^2}+(a_0+2a_1+a_2)\frac{z^2}{(1+z)^3}+\dots,$$

valid for small z, and then defines the first sum by means of the second. This is a definition of an entirely different type, since it is not linear in a_0, $a_1 z$, $a_2 z^2,\dots$; and the odd results to which it leads show that it is not a happy one. Thus Bromwich finds that $\sum 3^n z^n$ is summable inside the circle on $(-1, \frac{1}{2})$ as diameter,

but that $\sum 2^{-n}z^n$ is summable *outside* the circle on $(\frac{2}{3}, 2)$ as diameter, so that the region of summability does not include all the circle of convergence.

We add a remark concerning the calculations of Euler and Lacroix referred to in § 2.6. No Euler transform of $1-1!+2!-3!+\ldots$ is convergent, and it may seem remarkable that the method should have been applied to such a series successfully. We may, however, explain Euler's success as follows. If we write $a_n = (-1)^n n! = (-1)^n A_n$ and

$$A_n = \sum_p \alpha_{n,p}, \qquad \alpha_{n,0} = \int_0^2 e^{-t}t^n\,dt, \qquad \alpha_{n,p} = \int_{2^p+2^{p-1}-1}^{2^{p+1}+2^p-1} e^{-t}t^n\,dt \quad (p>0),$$

then it is easily verified that $\sum(-1)^n\alpha_{n,p}$ is summable $(\mathrm{E}, 2^{p+1}-1)$ to

$$\int_0^2 \frac{e^{-t}}{1+t}\,dt, \qquad \int_{2^p+2^{p-1}-1}^{2^{p+1}+2^p-1} \frac{e^{-t}}{1+t}\,dt$$

for $p = 0$ and $p > 0$ respectively; and if we add the results we obtain

$$\int \frac{e^{-t}}{1+t}\,dt,$$

in agreement with the B* sum found in § 8.11.

The remainders after $N+1$ terms in the appropriate Euler transforms are

$$2^{-N-1}\int_0^2 e^{-t}\frac{(1-t)^{N+1}}{1+t}\,dt, \qquad \frac{1}{2^{(p+1)(N+1)}}\int_{2^p+2^{p-1}-1}^{2^{p+1}+2^p-1} e^{-t}\frac{(2^{p+1}-1-t)^{N+1}}{1+t}\,dt,$$

which are $O(2^{-N-1})$, $O(e^{-2^p}2^{-N-1})$ respectively; and the Euler sums of the series are $O(e^{-2^p}2^p)$. Thus the error in taking only the first $N+1$ terms of the first P series, and ignoring the rest, is

$$O(2^{-N-1})+O(e^{-2^{P+1}}2^{-P-1}).$$

If, for example, we take $N = 10$, $P = 2$, we can easily prove that the error is less than $\cdot 001$.

This is naturally not quite what Euler does, and it would not be a convenient process; but Euler's process of reiterated transformation is roughly equivalent.

§ 8.5. Borel's earliest writings on divergent series, in his 'Mémoire sur les séries divergentes' [*AEN* (3), 16 (1899), 9–136] and the first edition of his book, contain a number of oversights corrected later by Hardy [*TCPS*, 19 (1903), 297–321, and *QJM*, 35 (1903), 22–66]. Here Hardy proves Theorems 122–7: these were rediscovered, with more concise proofs, by Perron, *MZ*, 6 (1920), 158–60 and 286–310. Sannia, *RP*, 42 (1917), 303–22, has extended the theorems in various directions.

Knopp, l.c. under § 3.7, observes that, since

$$e^{-x}\left(A_0x+A_1\frac{x^2}{2!}+\ldots\right) = \int_0^x e^{-(x-y)}e^{-y}(A_0+A_1y+\ldots)\,dy,$$

the B kernel of $0+a_0+a_1+\ldots$ is included in that of $a_0+a_1+\ldots$. He gives a further generalization in *RP*, 54 (1930), 331–4.

Theorem 128 is due to Knopp, l.c. under §§ 8.2–4.

§ 8.6. Hardy, l.c. under § 8.5, gives an example of a convergent series which is not absolutely summable.

§§ 8.7–8. The contents of these sections, except Theorem 130 and its corollary Theorem 131, are substantially Borel's. Theorem 130 was proved by Hardy, *MM*, 40 (1911), 161–5: the proof here is due to Landau, *AM*, 42 (1920), 95–8.

The region of (E, q) summability of $f(z)$ may be determined similarly. If ζ is the first singular point of $f(z)$ on a radius from O, and C_ζ is the circle

$$|q\zeta+z| < (q+1)|\zeta|,$$

then the region is the set of points interior to all C_ζ. It tends to the Borel polygon Π when $q \to \infty$. In particular (1) the series is summable (E, q), for some q, at all points inside Π, and (2) it is summable (E, q), for all $q > 0$, at any regular point on the circle of convergence. For all this see Knopp, l.c., and Rademacher, *Sitzungsberichte d. Berliner Math. Ges.* 21 (1922), 16–24.

Perron, *MZ*, 18 (1923), 157–72, has generalized Euler's method as follows. If

$$\gamma_m \geqslant 0, \qquad \sum \gamma_m = 1, \qquad F(z) = \sum \gamma_m z^{m+1},$$

$\sum b_m z^{m+1}$ is the result of developing $\sum a_n\{F(z)\}^{n+1}$ in powers of z, and $\sum b_m$ converges to A, then we say that $\sum a_n = A(F)$. The method is regular, and succeeds at every regular point on the circle of convergence. The (E, q) method corresponds to $F(z) = z/(q+1-qz)$.

§ 8.9. Hardy, *MM*, 43 (1913), 22–4.

§ 8.10. The three most elegant representations of $(1-z)^{-1}$ in its Mittag-Leffler star, viz.

$$(a)\ \lim_{\zeta\to 1} \sum \frac{\Gamma(1+\zeta n)}{\Gamma(1+n)} z^n, \qquad (b)\ 1+\lim_{\delta\to 0} \sum_1^\infty e^{-\delta n \log n} z^n, \qquad (c)\ \lim_{\delta\to 0} \sum \frac{z^n}{\Gamma(1+\delta n)},$$

all lead, after Theorem 135, to representations of a general $f(z)$. The first is due to Le Roy, *AT* (2), 2 (1900), 317–430 (323), and the second to Lindelöf, *J. de M.* (5), 9 (1903), 213–21. The third is mentioned by Mittag-Leffler in his address to the fourth international congress (Rome, 1908: see *Atti del IV Congresso Internaz.* i. 67–85). In his series of memoirs published in *AM* between 1899 and 1904 he uses the representation

$$(d)\quad \frac{1}{1-z} = \int e^{-t} \sum \frac{(t^\alpha z)^n}{\Gamma(1+\alpha n)}\, dt = \int e^{-t} E_\alpha(t^\alpha z)\, dt,$$

which is valid in the open region containing the origin and bounded by the curve

$$r = \left(\sec\frac{\theta}{\alpha}\right)^\alpha,$$

where $-\frac{1}{2}\alpha\pi < \theta < \frac{1}{2}\alpha\pi$ if $0 < \alpha \leqslant 2$ and $-\pi \leqslant \theta \leqslant \pi$ if $\alpha > 2$. The proof depends on the asymptotic properties of $E_\alpha(z)$. The region tends to the star if $\alpha \to 0$, but Mittag-Leffler does not, in these papers, give any simple formula valid throughout the star.

The behaviour of Mittag-Leffler's function

$$f(z) = E_\alpha(z) = \sum \frac{z^n}{\Gamma(\alpha n+1)},$$

for large $z = re^{i\theta}$, where $-\pi < \theta \leqslant \pi$, may be determined as follows. We have

$$\frac{1}{\Gamma(\alpha n+1)} = \frac{1}{2\pi i} \int e^u u^{-\alpha n-1}\, du,$$

where $u = \rho e^{i\phi}$, $u^{-\alpha n} = e^{-\alpha n \log u}$, $\log u$ has its principal value, and the contour C leaves the origin on its left and goes to infinity at each end in the left-hand half

of the plane. We consider two positions of C. In the first, C_0, it is formed by parts of the circle $\rho = 1$ and the radii $|\phi| = \frac{1}{2}\pi+\delta > \frac{1}{2}\pi$. In the second, C_1, it is everywhere at a distance from the origin greater than $(2r)^{1/\alpha}$, so that all zeros of $u^\alpha - z$ lie to its left. Then

$$f(z) = \frac{1}{2\pi i}\int_{C_1} \frac{e^u}{u}\sum\left(\frac{z}{u^\alpha}\right)^n du = \frac{1}{2\pi i}\int_{C_1} \frac{e^u u^{\alpha-1}}{u^\alpha - z}\,du;$$

and hence, by Cauchy's theorem,

$$f(z) = \frac{1}{2\pi i}\int_{C_0} \frac{e^u u^{\alpha-1}}{u^\alpha - z}\,du + \sum R_k = I(z) + \sum R_k,$$

where the R_k are the residues of the integrand at the poles, if any, which lie between C_0 and C_1. If δ is chosen so that no pole lies on C_0, and is sufficiently small, then these poles are given by $u = r^{1/\alpha}e^{i(\theta+2k\pi)/\alpha}$, where k runs through integers satisfying

(1) $$-\tfrac{1}{2}\alpha\pi \leqslant \theta + 2k\pi \leqslant \tfrac{1}{2}\alpha\pi;$$

and $$\alpha R_k = \exp\{r^{1/\alpha}e^{i(\theta+2k\pi)/\alpha}\}$$

is one of the values of $e^{z^{1/\alpha}}$. Now

$$I(z) = -\frac{1}{2\pi i}\int_{C_0}\left\{\frac{1}{z}+\frac{u^\alpha}{z^2}+\ldots+\frac{u^{(p-1)\alpha}}{z^p}-\frac{u^{p\alpha}}{z^p(u^\alpha - z)}\right\}u^{\alpha-1}e^u\,du$$

$$= -\sum_{m=1}^{p}\frac{z^{-m}}{\Gamma(1-m\alpha)} + R_p(z),$$

where $$R_p(z) = \frac{1}{2\pi i z^p}\int_{C_0}\frac{u^{p\alpha+\alpha-1}e^u}{u^\alpha - z}\,du = O\left(\frac{1}{|z|^{p+1}}\right)$$

for large z. Thus $I(z)$ has the asymptotic expansion

(2) $$I(z) \sim -\frac{z^{-1}}{\Gamma(1-\alpha)} - \frac{z^{-2}}{\Gamma(1-2\alpha)} - \ldots.$$

We must now distinguish the cases $\alpha < 2$, $\alpha > 2$. If $\alpha = 2$ then $f(z) = \cosh\sqrt{z}$, and tends to infinity in any direction except that of the negative real axis.

(1) If $0 < \alpha < 2$ and $\frac{1}{2}\alpha\pi < |\theta| \leqslant \pi$, then there is no k satisfying (1). In this case $E_\alpha(z)$ has the asymptotic expansion (2). If, on the other hand, $|\theta| \leqslant \frac{1}{2}\alpha\pi$, then we have to take account of the exponential term R_0, and

(3) $$\alpha E_\alpha(z) \sim e^{z^{1/\alpha}}$$

when $z \to \infty$ in the angle $|\theta| \leqslant \frac{1}{2}\alpha\pi$.

(2) If $\alpha > 2$ then we have always to take account of at least one exponential term. The modulus of such a term is

$$\frac{1}{\alpha}\exp\left(r^{1/\alpha}\cos\frac{\theta+2k\pi}{\alpha}\right),$$

and this is larger for $k = 0$ than for any other relevant k, except when $\theta = \pi$. Thus (3) holds for all θ except $\theta = \pi$. When $\theta = \pi$ there are two terms of equal importance, with $k = 0$ and $k = -1$. These combine to give

(4) $$\frac{1}{\alpha}\{\exp(r^{1/\alpha}e^{i\pi/\alpha}) + \exp(r^{1/\alpha}e^{-i\pi/\alpha})\} = \frac{2}{\alpha}\exp\left(r^{1/\alpha}\cos\frac{\pi}{\alpha}\right)\cos\left(r^{1/\alpha}\sin\frac{\pi}{\alpha}\right),$$

and $E_\alpha(z)$ behaves approximately like this function.

Wiman, *AM*, 29 (1905), 217–34, has shown that all the zeros of $E_\alpha(z)$ are real when $\alpha \geqslant 2$, and all but a finite number in any case.

We may also use the integral representation of $E_\alpha(z)$ to prove that, as was stated in § 4.11, $\sum z^n$ is summable (M) to $(1-z)^{-1}$ throughout its Mittag-Leffler star. For if $0 < |\theta| \leqslant \pi$, and α is sufficiently small, then $|\theta+2k\pi| > \frac{1}{2}\alpha\pi$ for all integral k. Hence

$$E_\alpha(z)-\frac{1}{1-z} = \frac{1}{2\pi i}\int \frac{e^u}{u}\left(\frac{u^\alpha}{u^\alpha-z}-\frac{1}{1-z}\right) du = \frac{1}{2\pi i}\,\frac{z}{1-z}\int \frac{e^u}{u}\,\frac{1-u^\alpha}{u^\alpha-z}\,du$$

round C_0, and this tends to 0 with α for all z of the star, and uniformly in any closed and bounded region interior to the star.

§ 8.11. Watson, *PTRS* (A), 211 (1912), 279–313. See also F. Nevanlinna, *ASF*, 12 (1916), no. 3, and Carleman, *Les fonctions quasi-analytiques* (Paris, 1926), ch. 5. The 'Phragmén-Lindelöf' theorems required will be found in Titchmarsh, *Theory of functions*, 176 et seq.

IX

THE METHODS OF EULER AND BOREL (2)

9.1. Some elementary lemmas. In this chapter we shall be concerned primarily with Tauberian theorems for Borel and Euler summability. We begin with three elementary theorems concerning the exponential and binomial series, on which much of our later work will depend. Their content is familiar.

THEOREM 137. *Suppose that $x > 0$ and*

(9.1.1) $$u_m = u_m(x) = e^{-x}\frac{x^m}{m!} \quad (m = 0, 1, 2,...),$$

so that $\sum u_m = 1$. Then

(1) *the largest u_m is u_M, where*

(9.1.2) $$M = [x],$$

two terms, u_{M-1} and u_M, being equal when x is an integer;

(2) *if*

(9.1.3) $$m = M+h$$

and $0 < \delta < 1$, then

(9.1.4) $$\sum_{|h|>\delta x} u_m = O(e^{-\gamma x}),$$

where $\gamma = \frac{1}{3}\delta^2$;

(3) *if*

(9.1.5) $$\tfrac{1}{2} < \zeta < \tfrac{2}{3},$$

then

(9.1.6) $$\sum_{|h|>x^\zeta} u_m = O(e^{-x^\eta}),$$

where η is any number less than $2\zeta-1$;

(4) *if $\lambda > 0$, then*

(9.1.7) $$\sum_{|h|>\lambda x^{\frac{1}{2}}} u_m < \epsilon,$$

for $x > x_0(\epsilon)$, $\lambda > \lambda_0(\epsilon)$;

(5) *if $|h| \leqslant x^\zeta$ then*

(9.1.8) $$u_m = \sqrt{\left(\frac{c}{\pi M}\right)} e^{-ch^2/M}\left\{1+O\left(\frac{|h|+1}{x}\right)+O\left(\frac{|h|^3}{x^2}\right)\right\},$$

where

(9.1.9) $$c = \tfrac{1}{2};$$

(6) *the estimates* (9.1.4) *and* (9.1.6) *remain valid if* u_m *is multiplied by any fixed power of* m.

THEOREM 138. *Suppose that* $q > 0$ *and*

$$u_m = u_m(n) = (q+1)^{-n}\binom{n}{m}q^{n-m} \quad (0 \leqslant m \leqslant n), \tag{9.1.10}$$

so that $\sum u_m = (q+1)^{-n}\sum\binom{n}{m}q^{n-m} = 1$. *Then*

(1) *the largest* u_m *is* u_M, *where*

$$M = \left[\frac{n+1}{q+1}\right], \tag{9.1.11}$$

two terms, u_{M-1} *and* u_M, *being equal when* $(n+1)/(q+1)$ *is an integer;*

(2) *clauses* (2)–(6) *of Theorem* 137 *hold with*

$$c = \frac{q+1}{2q}, \tag{9.1.12}$$

some positive $\gamma = \gamma(q, \delta)$ *in clause* (2),† *and* n *in the place of* x.

THEOREM 139. *Suppose that* $0 < k < 1$ *and*

$$u_m = u_m(n) = k^{n+1}\binom{m}{n}(1-k)^{m-n} \quad (m \geqslant n), \tag{9.1.13}$$

so that

$$\sum u_m = k^{n+1}\left\{1+(n+1)(1-k)+\frac{(n+1)(n+2)}{2!}(1-k)^2+\ldots\right\} = 1.$$

Then (1) *the largest* u_m *is* u_M, *where*

$$M = [n/k], \tag{9.1.14}$$

two terms, u_{M-1} *and* u_M, *being equal if* n/k *is an integer;*

(2) *clauses* (2)–(6) *of Theorem* 137 *hold with*

$$c = \frac{k}{2(1-k)}, \tag{9.1.15}$$

some positive $\gamma = \gamma(k, \delta)$ *in clause* (2), *and* n *in the place of* x.

We shall prove Theorems 137 and 139, which are the most important for our present purposes. The proof of Theorem 138 is like that of Theorem 139, only a little simpler.

9.2. Proof of Theorem 137. The first clause of the theorem is obvious because $u_m/u_{m-1} = x/m$.

† We do not assign a definite value to γ in this case.

Next, we divide $\sum u_m$ or $\sum u_{M+h}$ into the five pieces

(9.2.1)
$$\sum_{-M\leqslant h<-\delta x}+\sum_{-\delta x\leqslant h<-x^\zeta}+\sum_{|h|\leqslant x^\zeta}+\sum_{x^\zeta<h\leqslant \delta x}+\sum_{h>\delta x}$$
$$=\sum_{m=0}^{M_1-1}+\sum_{m=M_1}^{M_2-1}+\sum_{m=M_2}^{M_3}+\sum_{m=M_3+1}^{M_4}+\sum_{m=M_4+1}^{\infty}=S_1+S_2+S_3+S_4+S_5$$

(x being large enough to make $x^\zeta<\delta x$). Then

(9.2.2)
$$M_1=[x]-[\delta x],\quad M_2=[x]-[x^\zeta],\quad M_3=[x]+[x^\zeta],\quad M_4=[x]+[\delta x].$$

It is plain that

(9.2.3) $$S_1=O(xu_{M_1}),\qquad S_2=O(xu_{M_2}),\qquad S_4=O(xu_{M_3}).$$

Also $M_4+2>x+\delta x$, and so

(9.2.4)
$$S_5=u_{M_4+1}\left\{1+\frac{x}{M_4+2}+\frac{x^2}{(M_4+2)(M_4+3)}+\dots\right\}$$
$$<u_{M_4}\left\{1+\frac{1}{(1+\delta)}+\frac{1}{(1+\delta)^2}+\dots\right\}=O(u_{M_4}).$$

Hence, in order to prove clause (2) of the theorem, it is enough to prove that

(9.2.5) $$u_{M_1}=O(e^{-\gamma x}),\qquad u_{M_4}=O(e^{-\gamma x}).$$

Now
$$u_{M_1}=e^{-x}\frac{x^{M_1}}{M_1!}<e^{-x+M_1}\left(\frac{x}{M_1}\right)^{M_1}$$

and
$$x-\delta x-1<M_1=[x]-[\delta x]<x-\delta x+1.$$

Hence

(9.2.6)
$$u_{M_1}=O\left\{e^{-\delta x}\left(\frac{x}{x-\delta x-1}\right)^{x-\delta x+1}\right\}=O\left\{e^{-\delta x}\left(\frac{1}{1-\delta}\right)^{x-\delta x}\right\}=O(e^{-\Delta x}),$$

where

(9.2.7) $$\Delta=\delta-(1-\delta)\log\frac{1}{1-\delta}=\frac{\delta^2}{1.2}+\frac{\delta^3}{2.3}+\frac{\delta^4}{3.4}+\dots>\tfrac{1}{2}\delta^2.$$

Similarly

(9.2.8) $$u_{M_4}=O\left\{e^{\delta x}\left(\frac{1}{1+\delta}\right)^{x+\delta x}\right\}=O(e^{-\Delta' x}),$$

where

(9.2.9) $$\Delta'=-\delta+(1+\delta)\log(1+\delta)=\frac{\delta^2}{1.2}-\frac{\delta^3}{2.3}+\frac{\delta^4}{3.4}-\dots>\tfrac{1}{3}\delta^2.$$

It follows from (9.2.6)–(9.2.9) that (9.2.5) is true, with $\gamma=\frac{1}{3}\delta^2$. This proves (9.1.4).

We could prove (9.1.6) similarly, but it is shorter to base it on (9.1.8), which we have to prove in any case. For this, we write

$$x=M+f\quad(0\leqslant f<1),\qquad m=M+h,$$
$$\log u_m=-x+(M+h)\log x-\log\Gamma(M+h+1),$$

and approximate to $\log\Gamma(M+h+1)$ by the formula

$$\log\Gamma(y+1) = (y+\tfrac{1}{2})\log y - y + \tfrac{1}{2}\log 2\pi + O(y^{-1}).$$

A simple calculation gives

$$\log u_m = -\tfrac{1}{2}\log 2\pi M - \frac{h^2}{2M} + O\left(\frac{|h|+1}{x}\right) + O\left(\frac{|h|^3}{x^2}\right).\dagger$$

Since $|h| = o(x)$, $|h|^3 = o(x^2)$, this is equivalent to (9.1.8). Also, since

$$M^{-\frac{1}{2}} - x^{-\frac{1}{2}} = O(x^{-\frac{3}{2}}), \qquad \exp\left(\frac{h^2}{2M} - \frac{h^2}{2x}\right) = \exp\left\{O\left(\frac{h^2}{x^2}\right)\right\} = 1 + O\left(\frac{h^2}{x^2}\right),$$

we can replace M by x in (9.1.8) if we prefer to.

It is plain from (9.1.8) that u_{M_2} and u_{M_3} are $O(e^{-x^\eta})$: η may be any number less than $2\zeta-1$. It follows that S_2 and S_4 are $O(e^{-x^\eta})$; and this proves clause (3) of the theorem.

As regards clause (4), the sum is

$$O\left\{x^{-\frac{1}{2}} \sum_{\lambda\sqrt{x}<h\leqslant x\zeta} e^{-h^2/2x}\left(1+\frac{h}{x}+\frac{h^3}{x^2}\right)\right\} + O(e^{-x^\eta}),$$

and the first term here is

$$O\left\{x^{-\frac{1}{2}} \int_{\lambda\sqrt{x}}^{\infty} e^{-t^2/2x}\,dt + x^{-\frac{3}{2}} \int_{\lambda\sqrt{x}}^{\infty} e^{-t^2/2x} t\,dt + x^{-\frac{5}{2}} \int_{\lambda\sqrt{x}}^{\infty} e^{-t^2/2x} t^3\,dt\right\}.$$

The last two terms are $O(x^{-\frac{1}{2}})$, and the first is

$$O\left(\int_{\lambda}^{\infty} e^{-\frac{1}{2}w^2}\,dw\right)$$

which is small for large λ.

There remains clause (6). It is plain that, when u_m is replaced by $m^K u_m$, our estimates of S_1, S_2, and S_4 are affected only by a power of x, and that these sums are still of the orders required. The same is true of u_{M_4}; and

$$\begin{aligned} S_5 &= (M_4+1)^K u_{M_4+1} + (M_4+2)^K u_{M_4+2} + \dots \\ &\leqslant (2M_4)^K u_{M_4+1}\left\{1^K + \frac{2^K x}{M_4+2} + \frac{3^K x^2}{(M_4+2)(M_4+3)} + \dots\right\} \\ &\leqslant (2M_4)^K u_{M_4}\left\{1^K + \frac{2^K}{1+\delta} + \frac{3^K}{(1+\delta)^2} + \dots\right\} = O(x^K u_{M_4}). \end{aligned}$$

Thus S_5 also is of the order required, and this completes the proof.

† We must write $|h|+1$, not $|h|$, in the first error term, in order to include $h = 0$.

9.3. Proof of Theorem 139. The first clause of the theorem follows from
$$\frac{u_m}{u_{m-1}} = \frac{m}{m-n}(1-k).$$
The remainder of the proof follows the same lines as that of Theorem 137, though the calculations required are a little more elaborate.

We divide $\sum u_m$ again into five pieces: $M_1,\ldots$ are now
$$M_1 = \left[\frac{n}{k}\right]-[\delta n], \qquad M_2 = \left[\frac{n}{k}\right]-[n^\zeta], \qquad M_3 = \left[\frac{n}{k}\right]+[n^\zeta],$$
$$M_4 = \left[\frac{n}{k}\right]+[\delta n]:$$
we may suppose $\delta < k^{-1}-1$. As in § 9.2,
$$S_1 = O(nu_{M_1}), \qquad S_2 = O(nu_{M_2}), \qquad S_4 = O(nu_{M_3}).$$
Also $M_4+2 > n(k^{-1}+\delta)$, and so
$$\frac{u_m}{u_{m-1}} = \frac{m}{m-n}(1-k) < \frac{1+k\delta}{1-k+k\delta}(1-k) = 1-\frac{k^2\delta}{1-k+k\delta} = 1-\delta_1,$$
say, if $m > M_4+1$, and
$$S_5 < u_{M_4}\{1+(1-\delta_1)+(1-\delta_1)^2+\ldots\} = O(u_{M_4}).$$
Thus we have to prove, as in § 9.2, that u_{M_1} and u_{M_4} are $O(e^{-\gamma n})$ and that u_{M_2} and u_{M_3} are $O(e^{-n^\eta})$.

We have
$$l^{n+1}\sum_{m=n}^{\infty}\binom{m}{n}(1-l)^{m-n} = 1$$
for every l between 0 and 1, and so
$$u_m = k^{n+1}\binom{m}{n}(1-k)^{m-n} < \left(\frac{k}{l}\right)^{n+1}\left(\frac{1-k}{1-l}\right)^{m-n}.$$
If $m = M\pm[\delta n]$, this is $O(\theta^n)$, where
$$\theta = \theta(l) = \frac{k}{l}\left(\frac{1-k}{1-l}\right)^{1/k-1\pm\delta}.$$
Now $\theta(l)$ is 1 when $l = k$, tends to infinity when $l \to 0$ or $l \to 1$, and has a single minimum given by $l = k/(1\pm k\delta)$. Also
$$\theta'(k) = \pm\delta/(1-k) \neq 0.$$
It follows that $\theta(l)$ assumes values less than 1 for values of l on one side or the other of $l = k$ (according to the ambiguous sign). Hence we can certainly choose l so that $\theta < 1$, and thus u_{M_1} and u_{M_4} are $O(e^{-\gamma n})$.

This disposes of S_1 and S_4. We now suppose $|h| \leqslant n^{\zeta}$, write

$$M = \left[\frac{n}{k}\right] = \frac{n}{k} - f \quad (0 \leqslant f < 1), \qquad m = M+h,$$

$$\log u_m = (n+1)\log k + (m-n)\log(1-k) + \log\Gamma(m+1) - \log\Gamma(n+1) - \\ -\log\Gamma(m-n+1),$$

and approximate again by Stirling's theorem. We find that

$$\log u_m = -\tfrac{1}{2}\log 2\pi n + \log k - \tfrac{1}{2}\log(1-k) - \frac{h^2k^2}{2(1-k)n} + \\ + O\left(\frac{|h|+1}{n}\right) + O\left(\frac{|h|^3}{n^2}\right),$$

and this is equivalent to (9.1.8), with the c of (9.1.15) and the M of (9.1.14). The rest of the proof does not differ materially from that of Theorem 137.

9.4. Another elementary lemma. We shall also use another elementary lemma. Here sums without limits run over $-\infty < h < \infty$.

THEOREM 140. $$\sum e^{-ch^2/n} = \sqrt{\left(\frac{n\pi}{c}\right)} + O(1)$$

when $n \to \infty$, uniformly in any finite interval of positive c.

For the series is $1+2S$, where

$$S = \sum_1^{\infty} e^{-ch^2/n} = \int_1^{\infty} e^{-ct^2/n}\,dt + \sum_1^{\infty} \int_h^{h+1} (e^{-ch^2/n} - e^{-ct^2/n})\,dt$$

$$= \frac{1}{2}\sqrt{\left(\frac{n\pi}{c}\right)} - \int_0^1 e^{-ct^2/n}\,dt + T,$$

say. The integral here is less than 1. Also

$$e^{-ch^2/n} - e^{-ct^2/n} = \frac{2c}{n}\int_h^t ue^{-cu^2/n}\,du < \frac{4c}{n}(t-1)e^{-c(t-1)^2/n}$$

for $2 \leqslant h < t < h+1$, and is less than 1 in any case; so that

$$T < 1 + \frac{4c}{n}\int_1^{\infty} (t-1)e^{-c(t-1)^2/n}\,dt = 3.$$

We could obtain a much more precise result by using the formula

$$\sum e^{-ch^2/n} = \sqrt{\left(\frac{n\pi}{c}\right)} \sum e^{-\pi^2h^2n/c}.$$

9.5. Ostrowski's theorem on over-convergence. A famous theorem of Hadamard asserts that if $\sum b_m x^{\phi(m)}$ is an integral power series with a finite radius of convergence, and $\phi(m+1) > c\phi(m)$, where $c > 1$, then every point on the circle of convergence is a singular point of the function represented by the series. The condition on $\phi(m)$ has been generalized widely, Fabry having shown that it is sufficient to assume that $m^{-1}\phi(m) \to \infty$. Here, however, we are concerned with a generalization in a different direction due to Ostrowski, which Zygmund has shown to be deducible from theorems concerning Borel summability.

We say that a power series $\sum a_n z^n$ has a *gap* (n_k, n'_k) if $a_n = 0$ for $n_k < n < n'_k$.

THEOREM 141. *Suppose that* $\lambda > 0$. *Then there is a number* $\delta = \delta(\lambda) > 0$ *such that, if*

(i) $\sum a_n z^n$ *has an infinity of gaps* (n_k, n'_k) *for which*

$$n'_k/n_k \geqslant 1+\lambda > 1, \tag{9.5.1}$$

(ii) $A_n = a_0+a_1+\ldots+a_n = O\{(1+\delta)^n\}$,

(iii) $\sum a_n$ *is summable* (B) *to sum* A, *then*

$$A_{n_k} \to A.$$

Given λ, we can choose $\xi > 0$ so that

$$1-\xi > (1+\lambda)^{-\frac{1}{4}}, \qquad 1+\xi < (1+\lambda)^{\frac{1}{4}}, \tag{9.5.2}$$

and then δ so that

$$0 \leqslant \delta < \xi, \qquad \delta - \frac{(\xi \pm \delta)^2}{3(1+\delta)} < 0. \tag{9.5.3}$$

We write

$$e^{(1+\delta)x} = \left(\sum_{n<(1-\xi)x} + \sum_{(1-\xi)x}^{(1+\xi)x} + \sum_{n>(1+\xi)x}\right)\frac{(1+\delta)^n x^n}{n!} = e^x(P+Q+R),$$

where $P = P(\xi,\delta),\ldots$. If

$$y = (1+\delta)x, \qquad \eta = \frac{\xi+\delta}{1+\delta} < 1, \qquad (1-\eta)(1+\delta) = 1-\xi,$$

then
$$P(\xi,\delta) = e^{\delta x}e^{-x-\delta x}\sum_{n<(1-\xi)x}\frac{(1+\delta)^n x^n}{n!} = e^{\delta x}e^{-y}\sum_{n<(1-\eta)y}\frac{y^n}{n!}$$
$$= O(e^{\delta x - \frac{1}{3}\eta^2 y}) = O\left(\exp\left\{\delta x - \frac{(\xi+\delta)^2}{3(1+\delta)}x\right\}\right),$$

by Theorem 137 (9.1.4); and $R(\xi,\delta)$ has a similar bound with $\xi-\delta$ in

place of $\xi+\delta$. Hence $P \to 0$, $R \to 0$ if ξ and δ satisfy (9.5.3). In particular this is true when $\delta = 0$; and, since

$$P(\xi, 0)+Q(\xi, 0)+R(\xi, 0) = 1,$$

it follows that

(9.5.4) $$Q(\xi, 0) \to 1.$$

We now fix a ξ and a $\delta > 0$ satisfying (9.5.2) and (9.5.3), so that $\xi = \xi(\lambda)$, $\delta = \delta(\lambda)$, take

(9.5.5) $$x = \sqrt{(n_k n_k')},$$

write

$$e^{-x}A(x) = e^{-x}\sum A_n \frac{x^n}{n!} = e^{-x}\left\{\sum_{n<(1-\xi)x} + \sum_{(1-\xi)x}^{(1+\xi)x} + \sum_{n>(1+\xi)x}\right\} = P'+Q'+R',$$

and suppose that (ii) is true with our choice of δ. Then P' and R' are majorized by multiples of P and R, so that $P' \to 0$, $R' \to 0$; and therefore, by (iii),

(9.5.6) $$Q' \to \lim e^{-x}A(x) = A.$$

But

$$(1-\xi)x > \sqrt{\left(\frac{n_k n_k'}{1+\lambda}\right)} > n_k, \qquad (1+\xi)x < \sqrt{\{(1+\lambda)n_k n_k'\}} < n_k',$$

by (9.5.1), (9.5.2), and (9.5.5), so that every A_n in Q' is A_{n_k}. Thus (9.5.6) is $A_{n_k}Q(\xi, 0) \to A$, and therefore, by (9.5.4), $A_{n_k} \to A$.

We can easily deduce Ostrowski's and Hadamard's theorems.

THEOREM 142. *If $\sum a_n z^n$ has an infinity of gaps satisfying* (9.5.1), *and its sum $f(z)$ is regular at a point z_0 on the circle of convergence, then the partial sums $s_{n_k}(z)$ converge for $z = z_0$, and uniformly in a neighbourhood of $z = z_0$.*

We may suppose $z_0 = 1$. Then $A_n = O\{(1+\delta)^n\}$ for every positive δ. Also $\sum a_n z^n$ is summable (B) for $z = 1$, by Theorem 134, and uniformly summable in a neighbourhood of $z = 1$, so that (9.5.6) holds uniformly in the neighbourhood. Thus the conclusion follows from Theorem 141.

Hadamard's theorem is a corollary of Ostrowski's. We may suppose the radius of convergence to be 1, and it is enough to prove that $z = 1$ is a singular point. If we write the series as $\sum a_n z^n$, then $a_n = 0$ except when n is of the form $\phi(m)$, and every term whose coefficient is not 0 is the beginning and end of gaps satisfying (9.5.1). If $z = 1$ were a regular point, then the series would converge at points outside the circle of convergence; and therefore $z = 1$ is singular.

9.6. Tauberian theorems for Borel summability. We pass to the principal topic of the chapter. Our main purpose in this section and the next is to prove the theorem concerning Borel summability which corresponds to Tauber's Theorem 85, viz.

THEOREM 143. *If* $\sum a_n = A$ (B) *and*

$$a_n = o(n^{-\frac{1}{2}}), \tag{9.6.1}$$

then $\sum a_n$ *converges to* A.

Actually we shall prove a good deal more. Later (§ 9.13) we shall show that the o in (9.6.1) may be replaced by O; but this theorem (Theorem 156) is a good deal harder. All our conclusions will be true *a fortiori*, after Theorem 128, if $\sum a_n$ is summable (E, q). We need three lemmas.

THEOREM 144. *Suppose that*

$$\alpha \geqslant -1, \qquad 0 < \beta \leqslant 1, \qquad \rho > -1, \qquad 0 < H < 1; \tag{9.6.2}$$

that A_n^k *is defined as in* §§ 5.4–5;† *and that*

$$A_n^\alpha = o(n^\rho). \tag{9.6.3}$$

Then

$$A_n^{\alpha+\beta} = o(n^{\rho+\beta}) \tag{9.6.4}$$

and

$$A_m^{\alpha+\beta} - A_n^{\alpha+\beta} = o(|m-n|^\beta n^\rho) \tag{9.6.5}$$

uniformly for $0 < (1-H)n \leqslant m \leqslant (1+H)n$.

We have already proved (9.6.4) in § 5.7.‡ For (9.6.5), we use the formula

$$A_n^{\alpha+\beta} = \frac{1}{\Gamma(\beta)} \sum_{p=0}^{n} \frac{\Gamma(n-p+\beta)}{\Gamma(n-p+1)} A_p^\alpha . \text{§} \tag{9.6.6}$$

There are two cases, according as $m > n$ or $m < n$, and the proofs in these two cases differ only trivially. We take the case $m > n$. Then

$$\Gamma(\beta)(A_m^{\alpha+\beta} - A_n^{\alpha+\beta}) = \sum_{p=n+1}^{m} \frac{\Gamma(m-p+\beta)}{\Gamma(m-p+1)} A_p^\alpha +$$
$$+ \sum_{p=0}^{n} \left\{ \frac{\Gamma(m-p+\beta)}{\Gamma(m-p+1)} - \frac{\Gamma(n-p+\beta)}{\Gamma(n-p+1)} \right\} A_p^\alpha = S_1 + S_2,$$

say. Here, first,

$$S_1 = o\left\{ n^\rho \sum_{n+1}^{m} (m-p+1)^{\beta-1} \right\} = o\{(m-n)^\beta n^\rho\}.$$

† And $A_n^{-1} = a_n$, as in § 5.4.

‡ Strictly, for $\rho = \alpha$; but the proof for general ρ is substantially the same.

§ (9.6.6) is (5.4.8), extended, as in § 5.5, to general k and k'.

If $\beta = 1$, $S_2 = 0$. If $\beta < 1$, then the coefficient of A_p^α in S_2 is negative; and

$$S_2 = \sum_{p=0}^{n} \left\{\frac{\Gamma(n-p+\beta)}{\Gamma(n-p+1)} - \frac{\Gamma(m-p+\beta)}{\Gamma(m-p+1)}\right\} o(p^\rho) = S_3 + S_4,$$

where S_3 and S_4 extend over the ranges $0 \leqslant p \leqslant \frac{1}{2}n$ and $\frac{1}{2}n < p \leqslant n$ respectively.

In S_4 we may replace p^ρ by n^ρ, and then sum over the whole range $0 \leqslant p \leqslant n$. Thus

$$S_4 = o\left[n^\rho \sum_{p=0}^{n} \left\{\frac{\Gamma(n-p+\beta)}{\Gamma(n-p+1)} - \frac{\Gamma(m-p+\beta)}{\Gamma(m-p+1)}\right\}\right]$$

$$= o\left\{n^\rho \left(\sum_{q=0}^{n} - \sum_{q=m-n}^{m}\right) \frac{\Gamma(q+\beta)}{\Gamma(q+1)}\right\} = o\left\{n^\rho \left(\sum_{q=0}^{m-n-1} - \sum_{q=n+1}^{m}\right) \frac{\Gamma(q+\beta)}{\Gamma(q+1)}\right\}\dagger$$

$$= o\left\{n^\rho \sum_{q=0}^{m-n} \frac{\Gamma(q+\beta)}{\Gamma(q+1)}\right\} = o\left\{n^\rho \sum_{q=0}^{m-n} (q+1)^{\beta-1}\right\} = o\{(m-n)^\beta n^\rho\}.$$

Finally, in S_3

$$\frac{\Gamma(n-p+\beta)}{\Gamma(n-p+1)} - \frac{\Gamma(m-p+\beta)}{\Gamma(m-p+1)} = (n-p)^{\beta-1} - (m-p)^{\beta-1} + O(n^{\beta-2})$$

$$= O\{(m-n)n^{\beta-2}\} + O(n^{\beta-2}) = O\{(m-n)n^{\beta-2}\}.$$

Hence

$$S_3 = o\left\{(m-n)n^{\beta-2} \sum_{p \leqslant \frac{1}{2}n} p^\rho\right\} = o\{(m-n)n^{\beta+\rho-1}\}$$

$$= o\left\{\left(\frac{m-n}{n}\right)^{1-\beta} (m-n)^\beta n^\rho\right\} = o\{(m-n)^\beta n^\rho\}.$$

THEOREM 145. *If $k > 0$, $x \to \infty$, then* $e^{-x} \sum \dfrac{x^{n+k}}{\Gamma(n+k+1)} \to 1$.

For the sum is

$$\frac{e^{-x}}{\Gamma(k)} \sum \frac{1}{n!} \int_0^x (x-t)^{k-1} t^n \, dt = \frac{e^{-x}}{\Gamma(k)} \int_0^x (x-t)^{k-1} e^t \, dt = \frac{1}{\Gamma(k)} \int_0^x u^{k-1} e^{-u} \, du \to 1.$$

THEOREM 146. *If $k > 0$, and $\sum a_n \dfrac{t^n}{n!}$ is convergent for all t, then*

$$(9.6.7) \quad e^{-x} \sum A_n^k \frac{x^{n+k}}{\Gamma(n+k+1)} = \frac{1}{\Gamma(k)} \int_0^x (x-t)^{k-1} e^{-t} \sum A_n \frac{t^n}{n!} \, dt.$$

† Since $0 < m-n \leqslant Hn < n$.

If we write $x-u$ for t and multiply by e^x, the identity becomes

$$(9.6.8)\qquad \sum A_n^k \frac{x^{n+k}}{\Gamma(n+k+1)} = \frac{1}{\Gamma(k)} \int_0^x u^{k-1} e^u \sum A_n \frac{(x-u)^n}{n!}\, du.$$

The left-hand side of (9.6.8) is

$$\sum_{n=0}^{\infty} \frac{x^{n+k}}{\Gamma(n+k+1)} \sum_{p=0}^{n} \frac{\Gamma(n-p+k)}{\Gamma(k)\Gamma(n-p+1)} A_p$$

$$= \frac{x^k}{\Gamma(k)} \sum_{p=0}^{\infty} A_p \sum_{n=p}^{\infty} \frac{\Gamma(n-p+k)}{\Gamma(n+k+1)\Gamma(n-p+1)} x^n,$$

while the coefficient of A_p on the right-hand side is

$$\frac{1}{\Gamma(k)\Gamma(p+1)} \int_0^x u^{k-1}(x-u)^p e^u\, du = \frac{1}{\Gamma(k)\Gamma(p+1)} \sum_{q=0}^{\infty} \frac{1}{q!} \int_0^x u^{k+q-1}(x-u)^p\, du$$

$$= \frac{1}{\Gamma(k)} \sum_{q=0}^{\infty} \frac{\Gamma(q+k)}{\Gamma(q+p+k+1)\Gamma(q+1)} x^{q+p+k}$$

$$= \frac{x^k}{\Gamma(k)} \sum_{n=p}^{\infty} \frac{\Gamma(n-p+k)}{\Gamma(n+k+1)\Gamma(n-p+1)} x^n.$$

The identity may be regarded from another point of view whose full significance will appear in Ch. XI. Since, generally,

$$e^{-x} \sum u_n \frac{x^n}{n!} = \sum (-1)^n \Delta^n u_0 \frac{x^n}{n!},$$

the identity is

$$x^k \sum (-1)^n \Delta^n B_0 \frac{x^n}{n!} = \frac{1}{\Gamma(k)} \int_0^x (x-t)^{k-1} \left\{ \sum (-1)^n \Delta^n A_0 \frac{t^n}{n!} \right\} dt,$$

where
$$B_n = \frac{\Gamma(n+1)}{\Gamma(n+1+k)} A_n^k = \frac{C_n^k}{\Gamma(k+1)},$$

and C_n^k is the k-th Cesàro mean of $\sum a_n$; and this is

$$\sum (-1)^n \frac{\Gamma(n+k+1)}{\Gamma(n+1)} \Delta^n B_0 \frac{x^{n+k}}{\Gamma(n+k+1)} = \sum (-1)^n \Delta^n A_0 \frac{x^{n+k}}{\Gamma(n+k+1)}.$$

Thus the identity is equivalent to

$$(9.6.9)\qquad \Delta^n C_0^k = \frac{\Gamma(k+1)\Gamma(n+1)}{\Gamma(n+k+1)} \Delta^n A_0.$$

9.7. Tauberian theorems (*continued*). We can now prove

THEOREM 147. *If* $\rho \geqslant -\frac{1}{2}$,

$$(9.7.1)\qquad a_n = o(n^\rho),$$

and $\sum a_n$ *is summable* (B) *to* A, *then* $\sum a_n$ *is summable* (C, $2\rho+1$) *to* A.

The theorem shows that a series which is summable (B) and of finite order (i.e. $a_n = O(n^K)$ for some K) is necessarily summable (C, k) for sufficiently large k. Theorem 143 is the case $\rho = -\frac{1}{2}$.

We assume, first, that

(9.7.2) $$A_n^{k-\beta} = o(n^\rho),$$

where $\quad k \geqslant 0, \qquad 0 < \beta \leqslant 1, \qquad \rho > -1, \qquad \rho+\beta > 0$:

this reduces to (9.7.1) for $k = 0$, $\beta = 1$. By Theorem 144,

(9.7.3) $$A_m^k = o(n^{\rho+\beta}) \quad (m \leqslant n), \qquad A_m^k = o(m^{\rho+\beta}) \quad (m \geqslant n),$$

and

(9.7.4) $$A_m^k - A_n^k = o(|m-n|^\beta n^\rho) \quad (|m-n| < Hn).$$

We may take $A = 0$. Then, by Theorem 146, with $x = n$ and m for n,

(9.7.5) $$S = e^{-n} \sum A_m^k \frac{n^m}{\Gamma(m+k+1)} = o\left\{n^{-k} \int\limits_0^n (n-t)^{k-1}\, dt\right\} = o(1).$$

We write

(9.7.6)
$$S = e^{-n} A_n^k \sum \frac{n^m}{\Gamma(m+k+1)} + e^{-n} \sum (A_m^k - A_n^k) \frac{n^m}{\Gamma(m+k+1)} = S_1 + S_2,$$

(9.7.7) $$S_2 = e^{-n}\Bigl(\sum_{m<n-n^\zeta} + \sum_{n-n^\zeta}^{n+n^\zeta} + \sum_{m>n+n^\zeta}\Bigr) = S_2^{(1)} + S_2^{(2)} + S_2^{(3)},$$

where $\frac{1}{2} < \zeta < \frac{2}{3}$. Then (9.7.5) is

(9.7.8) $$S_1 + S_2^{(1)} + S_2^{(2)} + S_2^{(3)} = o(1).$$

Here, first, $$S_1 = n^{-k} A_n^k \{1+o(1)\},$$

by Theorem 145. Secondly,

$$S_2^{(1)} = O\Bigl(e^{-n} n^{\rho+\beta} \sum_{m<n-n^\zeta} \frac{n^m}{m!}\Bigr) = O(e^{-n^\eta}),$$

where $\eta = \eta(\zeta) > 0$, by (9.7.3) and Theorem 137 (3). Thirdly, after (9.7.3) and Theorem 137, (3) and (6),

$$S_2^{(3)} = O\Bigl(e^{-n} \sum_{m>n+n^\zeta} m^{\rho+\beta} \frac{n^m}{\Gamma(m+k+1)}\Bigr)$$
$$= O\Bigl(e^{-n} \sum_{m>n+n^\zeta} m^{\rho+\beta-k} \frac{n^m}{\Gamma(m+1)}\Bigr) = O(e^{-n^\eta});$$

and it now follows from (9.7.8) that

(9.7.9) $$n^{-k} A_n^k \{1+o(1)\} + S_2^{(2)} = o(1).$$

It remains to discuss $S_2^{(2)}$. Here we use (9.7.4) and Theorem 137 (5), which give

$$S_2^{(2)} = o\left\{e^{-n}n^\rho \sum_{n-n^\zeta}^{n+n^\zeta} |m-n|^\beta \frac{n^m}{\Gamma(m+k+1)}\right\} = o\left(e^{-n}n^{\rho-k} \sum_{n-n^\zeta}^{n+n^\zeta} |m-n|^\beta \frac{n^m}{m!}\right)$$

$$= o\left(e^{-n}n^{\rho-k} \sum_{|r|\leqslant n^\zeta} |r|^\beta \frac{n^{n+r}}{(n+r)!}\right) = o\left(n^{\rho-k-\frac{1}{2}} \sum_{|r|\leqslant n^\zeta} |r|^\beta e^{-r^2/2n}\right)$$

$$= o\left(n^{\rho-k-\frac{1}{2}} \int_{-\infty}^{\infty} |t|^\beta e^{-t^2/2n}\, dt\right) = o(n^{\rho-k+\frac{1}{2}\beta}).$$

Hence, finally, (9.7.9) becomes

(9.7.10) $$n^{-k}A_n^k\{1+o(1)\}+o(n^{\rho-k+\frac{1}{2}\beta}) = o(1),$$

i.e.

(9.7.11) $$A_n^k = o(n^k)+o(n^{\rho+\frac{1}{2}\beta}).$$

Taking $k = 0$, $\beta = 1$, so that $a_n = o(n^\rho)$, in (9.7.2) and (9.7.11), we obtain

(9.7.12) $$A_n = o(1)+o(n^{\rho+\frac{1}{2}}) = o(n^{\rho+\frac{1}{2}}).$$

Next, we suppose that $\nu\beta = 2\rho+1$, where ν is an integer and $\beta \leqslant 1$, and prove that

(9.7.13) $$A_n^{r\beta} = o(n^{\rho+\frac{1}{2}+\frac{1}{2}r\beta})$$

for $0 \leqslant r \leqslant \nu$. First, (9.7.13) is true for $r = 0$, by (9.7.12). We assume it true for $r = s < \nu$, and use (9.7.2) and (9.7.11), with $k = (s+1)\beta$ and $\rho+\frac{1}{2}+\frac{1}{2}s\beta$ in place of ρ. Then (9.7.11) gives

$$A_n^{(s+1)\beta} = o\{n^{(s+1)\beta}\}+o\{n^{\rho+\frac{1}{2}+\frac{1}{2}(s+1)\beta}\} = o\{n^{\rho+\frac{1}{2}+\frac{1}{2}(s+1)\beta}\},$$

since $$\tfrac{1}{2}(s+1)\beta \leqslant \tfrac{1}{2}\nu\beta = \rho+\tfrac{1}{2};$$

and this is (9.7.13) with $r = s+1$.

Thus (9.7.13) is true generally. Finally, taking $r = \nu$, we find that

$$A_n^{2\rho+1} = o(n^{2\rho+1}),$$

i.e. that $\sum a_n$ is summable $(C, 2\rho+1)$ to 0.

We have stated the theorem for summability (B), but it is equally true under the slightly weaker hypothesis of summability (B′). For, if $\sum a_n$ is summable (B′), then $0+a_0+a_1+\ldots$ is summable (B), by Theorem 126, and the terms of this series are of the same order as those of the original series. Hence it is summable $(C, 2\rho+1)$, and, therefore, by Theorem 47, the original series is summable $(C, 2\rho+1)$.

Incidentally we have proved

THEOREM 148. *If $a_n = o(1)$ and $\sum a_n$ is summable* (B), *then $A_n = o(n^{\frac{1}{2}})$.*

We shall need this theorem later.

9.8. Examples of series not summable (B). Theorem 147 shows that no series of finite order can be summable (B) unless it is summable (C). On the other hand, we can find examples of series, summable (C) (and so of finite order) but not summable (B).

(i) The series $\sum n^{\alpha-1}e^{Ain^\alpha}$, where $0<\alpha<1$, $A>0$, is summable (C,k) for every positive k, by Theorem 84, but is not convergent. If $\alpha<\frac{1}{2}$, then the general term is $o(n^{-\frac{1}{2}})$; and therefore, by Theorem 143, the series is not summable (B). *A fortiori*, by Theorem 128, it is not summable (E,q) for any q.

(ii) If $\alpha=\frac{1}{2}$ then the general term is $O(n^{-\frac{1}{2}})$ but not $o(n^{-\frac{1}{2}})$. The conclusions of (i) are still correct but (in default of direct analysis) we must appeal to the more difficult Theorem 156, which we have yet to prove.

(iii) Suppose that $A_n=(-1)^m m$ when $n=m^2$ and $A_n=0$ otherwise: the series is then that derived from

$$-x(1-x)+2x^4(1-x)-3x^9(1-x)+\dots$$

by writing it as a power series and putting $x=1$. If $N^2\leqslant n<(N+1)^2$, then

$$C_n^1(A)=\frac{1}{n+1}\{-1+2-3+\dots+(-1)^N N\}=O\left(\frac{N}{n}\right)=O(n^{-\frac{1}{2}}),$$

so that the series is summable (C, 1) to 0; but it is not summable (B). In fact, if in

$$S(x)=e^{-x}\sum(-1)^m m\frac{x^{m^2}}{m^2!}$$

we take $x=N^2$ and $m=N+\mu$, then it is easy to show, by use of the approximations of Theorem 137, that $S(N^2)$ is dominated by the terms for which $|\mu|<N^\beta$, where $0<\beta<\frac{1}{6}$, and that

$$S(N^2)=\frac{(-1)^N}{\sqrt{(2\pi)}}\sum_{|\mu|<N^\beta}(-1)^\mu e^{-2\mu^2}+o(1)$$

assumes values of alternating sign, numerically greater than $\frac{1}{4}$, when $N\to\infty$.

9.9. A theorem in the opposite direction. The examples of the last section show that summability (C) does not necessarily involve summability (B), and *a fortiori* that it does not involve summability (E,q) for any q. It is natural to ask what strengthening of the hypothesis is necessary to secure such a conclusion, and the simplest theorem in this direction is as follows.

THEOREM 149. *If*

$$C_n^1(A)=\frac{A_0+A_1+\dots+A_n}{n+1}=A+o(n^{-\frac{1}{2}}), \tag{9.9.1}$$

then $\sum a_n$ is summable (E,q) to A for every positive q, and a fortiori *summable* (B).

We may take $A=0$, when, by (8.3.4), we have to show that

$$A_n^1=A_0+A_1+\dots+A_n=o(n^{\frac{1}{2}})$$

implies

$$A_n^{(q)}=\frac{1}{(q+1)^n}\sum_{m=0}^{n}\binom{n}{m}q^{n-m}A_m=\sum_{m=0}^{n}v_m A_m=o(1):$$

here $v_m = v_m(n)$. The largest v_m, say v_M, is given by (9.1.11), and, by Theorem 138 (2),

(9.9.2) $$v_m \leqslant v_M < Hn^{-\frac{1}{2}},$$

where H is independent of n and m. We choose μ so that

(9.9.3) $$|A_p+A_{p+1}+\ldots+A_r| < \epsilon r^{\frac{1}{2}}$$

for $r \geqslant p \geqslant \mu$, suppose n so large that $M > \mu$, and write

$$\mathrm{A}_n^{(q)} = \left(\sum_0^{\mu-1}+\sum_{\mu}^{M}+\sum_{M+1}^{n}\right)v_m A_m = S_1+S_2+S_3,$$

say. Then, since v_m decreases on either side of $m = M$, we have

$$|S_2| \leqslant v_M . \epsilon M^{\frac{1}{2}} \leqslant H\epsilon, \qquad |S_3| \leqslant v_M . \epsilon n^{\frac{1}{2}} \leqslant H\epsilon,$$

by (9.9.2) and (9.9.3); and $S_1 \to 0$ when μ is fixed and $n \to \infty$. Hence $|\mathrm{A}_n^q| \leqslant 3H\epsilon$ for sufficiently large n.

Theorem 149 is a best possible theorem in the sense that the o of (9.9.1) cannot be replaced by O. This is shown by the series of § 9.8 (iii): we have seen that in this case $C_n^1(A) = O(n^{-\frac{1}{2}})$, but that the series is not summable (B).

9.10. The (e, c) **method of summation.** Our primary object in the rest of this chapter is the proof of Theorem 156, the generalization of Theorem 143 in which o is replaced by O. The proof is rather difficult, and will be simplified considerably by a preliminary study of some other methods of summation. These methods appear here as auxiliaries, but they have also some independent interest.

We shall be concerned primarily with delicately divergent series, among which we distinguish three classes, the class $\mathfrak{P}$ for which

(9.10.1) $$a_n = o(1),$$

the class $\mathfrak{Q}$ for which

(9.10.2) $$A_n = o(n^{\frac{1}{2}}),$$

and the class $\mathfrak{R}$ for which

(9.10.3) $$a_n = O(n^{-\frac{1}{2}}).$$

It is plain that $\mathfrak{P}$ includes $\mathfrak{R}$, while $\mathfrak{Q}$ does not (and *a fortiori* does not include $\mathfrak{P}$). We shall, however, find that all series of $\mathfrak{P}$, summable by any of the methods which we are considering, belong to $\mathfrak{Q}$. We have seen already† that this is true for series summable (B).

We shall often use a number ζ which, as in §§ 9.1–3, lies between $\frac{1}{2}$ and $\frac{2}{3}$. We shall be dealing, as there, with sums with respect to h, or

† See the last remark of § 9.7.

integrals with respect to t, in which the contribution of the parts $|h| > n^{\zeta}$, or $|t| > x^{\zeta}$, is $O(e^{-n^{\eta}})$, or $O(e^{-x^{\eta}})$, where $\eta = \eta(\zeta) > 0$. These 'tails' of the sums or integrals will therefore be trivial, and we shall hold ourselves at liberty to reject or retain them at our convenience.

We define *summability* (e, c), in the first instance, by saying that

(9.10.4) $$\sum a_n = A \quad (e, c)$$

means

(9.10.5) $$\sqrt{\left(\frac{c}{\pi n}\right)} \sum e^{-ch^2/n} A_{n+h} \to A$$

when $n \to \infty$. Here $c > 0$, the variable of summation h runs from $-\infty$ to ∞, and A_m is to be interpreted as 0 when $m < 0$. It is, however, usually convenient to vary this definition by introducing a continuous parameter x. We define $A(t)$ as A_n when $n \leqslant t < n+1$, and take

(9.10.6) $$\sqrt{\left(\frac{c}{\pi x}\right)} \int e^{-ct^2/x} A(x+t)\, dt \to A,\dagger$$

when $x \to \infty$ continuously, as an alternative definition of summability.

We begin by proving

THEOREM 150. *If* $a_n = o(1)$, *and* $\sum a_n$ *is summable* (B), (E, q), *or* (e, c), *in accordance with either of the definitions* (9.10.5) *or* (9.10.6), *then* $A_n = o(n^{\frac{1}{2}})$.

We proved this for B in § 9.7:‡ *a fortiori*, after Theorem 128, it is true for (E, q). We have therefore only to prove it for series summable (e, c). We take the first definition, the proof for the second being similar.

After Theorem 140,

$$A_n \sqrt{\left(\frac{c}{\pi n}\right)} \sum e^{-ch^2/n} = A_n\{1+o(1)\},$$

and therefore, if (9.10.5) is true,

(9.10.7) $$\sqrt{\left(\frac{c}{\pi n}\right)} \sum (A_{n+h} - A_n) e^{-ch^2/n} = A + o(1) - A_n\{1+o(1)\}.$$

Since $A_n = o(n)$, $A_{n+h} = o(n+|h|)$, we may neglect the terms for which $|h| > n^{\zeta}$; and $A_{n+h} - A_n = o(|h|)$ in the remainder. Hence the left-hand side of (9.10.7) is

$$o\left(n^{-\frac{1}{2}} \sum |h| e^{-ch^2/n}\right) = o\left(n^{-\frac{1}{2}} \int |t| e^{-ct^2/n}\, dt\right) = o(n^{-\frac{1}{2}} . n) = o(n^{\frac{1}{2}}),$$

and therefore $A_n = o(n^{\frac{1}{2}})$.

† Here t runs from $-\infty$ to ∞.

‡ Theorem 148.

Next, we prove

THEOREM 151. *If $A_n = o(\sqrt{n})$, then summability* (B) *is equivalent to summability* $(e, \frac{1}{2})$ *by either of the definitions* (9.10.5) *and* (9.10.6).

We prove the implication $B \to (e, \frac{1}{2})$: our arguments will be plainly reversible. We may suppose $A = 0$.

(1) We prove first that

$$e^{-n} \sum A_m \frac{n^m}{m!} = o(1) \tag{9.10.8}$$

implies

$$\sum e^{-h^2/2n} A_{n+h} = o(\sqrt{n}) \tag{9.10.9}$$

when $n \to \infty$ by integral values. We may neglect the 'tails' of the sums, for which $m = n+h$, $|h| > n^{\zeta}$, and use the approximation (9.1.8). Then (9.10.8) is

$$\sum e^{-h^2/2n}\left\{1+O\left(\frac{|h|+1}{n}\right)+O\left(\frac{|h|^3}{n^2}\right)\right\}A_{n+h} = o(\sqrt{n}), \tag{9.10.10}$$

and $A_{n+h} = o(\sqrt{n})$. Hence the O terms in (9.10.10) give

$$o\left\{\frac{1}{\sqrt{n}} \int e^{-t^2/2n}(|t|+1)\, dt\right\}, \qquad o\left\{\frac{1}{n^{\frac{3}{2}}} \int e^{-t^2/2n}|t|^3\, dt\right\}$$

respectively. Since these are $o(\sqrt{n})$, we obtain (9.10.9).

(2) Next, we replace (9.10.9) by

$$\int e^{-t^2/2n} A(n+t)\, dt = o(\sqrt{n}). \tag{9.10.11}$$

The difference is

$$\sum \int_h^{h+1} (e^{-h^2/2n} - e^{-t^2/2n}) A(n+t)\, dt, \tag{9.10.12}$$

and here again we may neglect the tails and suppose $A(n+t) = o(\sqrt{n})$. Any part of (9.10.12) in which h is bounded is plainly $o(\sqrt{n})$, so that we may suppose h and t large. Then

$$e^{-h^2/2n} - e^{-t^2/2n} = O\left(\frac{|t|}{n} e^{-t^2/3n}\right),$$

and (9.10.12) gives

$$o\left(n^{-1} \int e^{-t^2/3n}|t|\, dt\right) = o(\sqrt{n}).$$

Thus we obtain (9.10.11).

(3) Finally, we replace the n of (9.10.11) by a continuous x. If $x = n+f$, where $0 < f < 1$, then the integral in (9.10.6), with $c = \frac{1}{2}$, is

$$\int e^{-t^2/2(n+f)} A(n+f+t)\, dt = \int e^{-(t-f)^2/2(n+f)} A(n+t)\, dt,$$

and what we have to prove is that

$$\int \{e^{-(t-f)^2/2(n+f)} - e^{-t^2/2n}\} A(n+t)\, dt = o(\sqrt{n}).$$

We may again neglect tails, replace $A(n+t)$ by $o(\sqrt{n})$, and suppose t large. We have

$$e^{-(t-f)^2/2(n+f)} - e^{-t^2/2n} = f\frac{d}{dw}\{e^{-(t-w)^2/2(n+w)}\}$$

for a w of $(0, 1)$, and this is

$$O\left(e^{-(t-w)^2/2(n+w)}\left\{\frac{|t-w|}{n+w} + \frac{(t-w)^2}{2(n+w)^2}\right\}\right) = O\left\{\left(\frac{|t|}{n} + \frac{t^2}{n^2}\right)e^{-t^2/3n}\right\}.$$

Hence our error is

$$o\left(\frac{1}{\sqrt{n}}\int e^{-t^2/3n}|t|\, dt\right) + o\left(\frac{1}{n^{\frac{3}{2}}}\int e^{-t^2/3n}|t|^2\, dt\right) = o(\sqrt{n}) + o(1) = o(\sqrt{n}),$$

and this completes the proof.

If $a_n = o(1)$, and the series is summable, then $A_n = o(\sqrt{n})$, by Theorem 150, and the result of Theorem 151 holds *a fortiori*. Throughout the rest of this chapter we shall suppose that $a_n = o(1)$, not troubling to ask whether this hypothesis is essential; but there is one place in Ch. XII (§ 12.15) where it will be essential to have proved Theorem 151 without it.

THEOREM 152. *If* $a_n = o(1)$, *and* $\sum a_n$ *is summable* (E, q), *then it is summable* (e, c), *where* $c = (q+1)/2q$, *to the same sum; and conversely.*

We write

$$M = \left[\frac{n+1}{q+1}\right], \qquad m = M+h, \qquad r = \frac{1}{q+1}, \qquad r' = \frac{q}{q+1};$$

and we may then replace the Euler mean of A_n by

$$\chi_n = \frac{1}{(q+1)^n} \sum_{|h|<n^{\zeta}} \binom{n}{M+h} q^{n-M-h} A_{M+h}.$$

Here, after Theorem 138,

$$\frac{1}{(q+1)^n}\binom{n}{M+h} q^{n-M-h} = \frac{1}{\sqrt{(2\pi rr'n)}} e^{-h^2/(2rr'n)}\left\{1 + O\left(\frac{|h|+1}{n}\right) + O\left(\frac{|h|^3}{n^2}\right)\right\},$$

and it follows, as in the proof of Theorem 151, that $\chi_n \to A$ is equivalent to

$$\frac{1}{\sqrt{(2\pi rr'n)}} \sum e^{-h^2/(2rr'n)} A_{M+h} \to A.$$

This is effectively (9.10.5), with $c = (q+1)/2q$ and M in the place of n.

Theorem 151, with $a_n = o(1)$, corresponds to $q = \infty$.

If $c = (q+1)/2q$ then c decreases from ∞ to $\frac{1}{2}$ when q increases from 0 to ∞. If we suppose $a_n = o(1)$, and combine Theorems 152 and 118, we find that *if $a_n = o(1)$ and $\frac{1}{2} < c' < c$, then summability (e, c) implies summability (e, c')*. We shall, however, see in § 9.11 that this theorem is incomplete, the result being true for $0 < c' < c$.

9.11. The circle method of summation. We shall say that

(9.11.1) $$\sum a_n = A \quad (\gamma, k),$$

where $0 < k < 1$, if (i) $\sum a_n x^n$ is convergent for $|x| < 1$, and (ii) the Taylor series for $f(1-k+ky)$, viz.

(9.11.2) $$\sum \frac{f^{(n)}(1-k)}{n!}(ky)^n = \sum b_n y^n$$

is convergent for $y = 1$.

Here $B_m = b_0+b_1+...+b_m$ is the coefficient of y^m in the expansion, for $|y| < 1$, of $(1-y)^{-1}f(1-k+ky)$. If $x = 1-k+ky$ then $1-x = k(1-y)$, and so

$$B_m = \left[\frac{k}{1-x}f(x)\right]_{y^m} = \left[\frac{k}{1-x}\sum a_n x^n\right]_{y^m}$$
$$= k[\sum A_n x^n]_{y^m} = k[\sum A_n(1-k+ky)^n]_{y^m}.$$

Since y^m occurs in $k(1-k+ky)^n$ only when $n \geqslant m$, and then with coefficient $\binom{n}{m}k^{m+1}(1-k)^{n-m}$, it follows that

(9.11.3)
$$B_m = k^{m+1}\left\{A_m+(m+1)(1-k)A_{m+1}+\frac{(m+1)(m+2)}{2!}(1-k)^2A_{m+2}+...\right\}.$$

It may be verified at once that the method is regular, and that the relation between b_m and a_n is

(9.11.4)
$$b_m = k^m\left\{a_m+(m+1)(1-k)a_{m+1}+\frac{(m+1)(m+2)}{2!}(1-k)^2a_{m+2}+...\right\}.$$

These formulae have been obtained under the hypothesis that the radius of convergence of $\sum a_n x^n$ is 1, but we might abandon this hypothesis and say simply that $\sum a_n$ is summable (γ, k), to sum A, if B_m, defined by (9.11.3), exists for all m and tends to A. It is, however, more convenient for our present purpose to keep the restriction, and the theorems which follow depend upon it.

THEOREM 153. *If $\sum a_n$ is summable (γ, k), and $0 < l < k$, then $\sum a_n$ is summable (γ, l) to the same sum.*

We write

$$x = 1-k+ky = 1-l+lz, \qquad y = 1-\frac{l}{k}+\frac{l}{k}z,$$

$$f(x) = \sum a_n x^n = \sum b_n y^n = \sum c_n z^n,$$

and denote the series $\sum a_n$, $\sum b_n$, and $\sum c_n$ by A, B, and C. If A is summable (γ, k) then B is convergent, and therefore summable $(\gamma, l/k)$. But this means that C is convergent, i.e. that A is summable (γ, l).

There is a theorem for the (γ, k) method corresponding to Theorems 151 and 152.

THEOREM 154. *If $a_n = o(1)$, and $\sum a_n$ is summable (γ, k), then it is summable (e, c), where*

(9.11.5) $$c = \frac{k}{2(1-k)},$$

to the same sum; and conversely.

The (γ, k) mean, of rank n, is

$$B_n = k^{n+1} \sum_{p=n}^{\infty} \binom{p}{n} (1-k)^{p-n} A_p = \sum_{h=-\infty}^{\infty} u_{m+h} A_{m+h},$$

where $$u_p = \binom{p}{n} k^{n+1} (1-k)^{p-n}$$

when $p \geqslant n$, $u_p = 0$ when $p < n$, and $m = [n/k]$. After Theorem 139, we may neglect the terms for which $|h| > n^{\zeta}$, and write

$$u_{m+h} = \sqrt{\left(\frac{c}{\pi m}\right)} e^{-ch^2/m} \left\{1+O\left(\frac{|h|+1}{m}\right)+O\left(\frac{|h|^3}{m^2}\right)\right\},$$

where c is given by (9.11.5), in the remainder. The result then follows as in the proofs of Theorems 151 and 152.

When k increases from 0 to 1, c increases from 0 to ∞. Hence Theorems 153 and 154 give

THEOREM 155. *If $a_n = o(1)$ and $\sum a_n$ is summable (e, c), then it is summable (e, c'), to the same sum, for $0 < c' < c$.*

This theorem will play an essential part in the proof of Theorem 156.

9.12. Further remarks on Theorems 150–5. We have supposed throughout §§ 9.10–11 (except in Theorems 151 and 153) that $\sum a_n$ is a series of the class $\mathfrak{P}$, but the proofs of Theorems 152, 154, and 155, like that of Theorem 151, require only that it should belong to $\mathfrak{Q}$: Theorem 150 shows that $\mathfrak{P}$ is, for our purposes here, a subclass of $\mathfrak{Q}$.

Theorems 151, 152, 154, and 155 are actually true for much wider classes of series. We need them only as tools for the proof of Theorem 156, and it is not necessary to consider how far they may be generalized; but what we have proved falls far short of the ultimate truth. Thus Hyslop has proved that the B and

$(e, \frac{1}{2})$ methods are equivalent whenever $a_n = O(n^K)$ for some K, and the scope of Theorems 152 and 154 might be extended similarly. It would follow that Theorem 155 also is true for all series of finite order.

The range of this theorem might be extended still farther by the use of different methods. Let us consider, for convenience, the analogue of the theorem for continuous limits. We say that $f(x) \to l\ (e, c)$ if

$$f_c(x) = \sqrt{\left(\frac{c}{\pi x}\right)} \int e^{-ct^2/x} f(x+t)\,dt \to l, \tag{9.12.1}$$

the integral being a Lebesgue integral. Then it may be proved that

$$f_b(x) = \sqrt{\left(\frac{ab}{\pi}\right)} \int\limits_0^{ax/b} \frac{(a-b)x}{(ax-bt)^{\frac{3}{2}}} \exp\left\{-\frac{ab(t-x)^2}{ax-bt}\right\} f_a(t)\,dt \tag{9.12.2}$$

whenever $0 < b < a$ and the integral (9.12.1) is convergent for $c = b$; and deduced that, subject only to this last condition,

$$f(x) \to l\ (e, a) \rightarrow f(x) \to l\ (e, b).$$

This is another illustration of the principle of § 4.12.

9.13. The principal Tauberian theorem. We are now in a position to prove our main theorem.

THEOREM 156. *If* $\sum a_n = A$ (B) *and*

$$a_n = O(n^{-\frac{1}{2}}), \tag{9.13.1}$$

then $\sum a_n$ *converges to* A.

We begin by proving that $A_n = O(1)$. It follows from Theorems 150 and 151 that

$$\sqrt{\left(\frac{c}{\pi n}\right)} \sum e^{-ch^2/n} A_{n+h} = A + o(1) \tag{9.13.2}$$

for $c = \frac{1}{2}$. By Theorem 140

$$\sqrt{\left(\frac{c}{\pi n}\right)} \sum e^{-ch^2/n} = 1 + o(1),$$

and therefore

$$\begin{aligned} A_n\{1+o(1)\} &= \sqrt{\left(\frac{c}{\pi n}\right)} A_n \sum e^{-ch^2/n} \\ &= \sqrt{\left(\frac{c}{\pi n}\right)} \sum e^{-ch^2/n}(A_n - A_{n+h}) + A + o(1). \end{aligned} \tag{9.13.3}$$

We may restrict the summation to $|h| < n^{\zeta}$. Then $A_{n+h} - A_n$ is $O(n^{-\frac{1}{2}}|h|)$, and the sum is

$$O\left(n^{-1} \sum_{|h| < n^{\zeta}} e^{-ch^2/n} |h|\right) = O\left(n^{-1} \int e^{-ct^2/n} |t|\,dt\right) = O(1).$$

Hence $A_n\{1+o(1)\} = O(1)$, i.e. $A_n = O(1)$.

The next (and the critical) stage of the proof depends upon 'Vitali's theorem'. Suppose that D is an open and connected region of z, that $\phi_n(z)$ is, for each n, an analytic function of z regular in D, that $\phi_n(z)$ is uniformly bounded in D, and that $\phi_n(z_k)$ tends to a limit $\phi(z_k)$, for each of a set of points z_k with at least one limit point in D. Then there is an analytic function $\phi(z)$ such that $\phi_n(z) \to \phi(z)$ uniformly in any closed and bounded region R contained in D.

To apply the theorem, we suppose $c = \gamma + i\delta$ complex; choose $\delta_0, \gamma_0, \gamma_1$ so that $\delta_0 > 0$, $0 < \gamma_0 < \frac{1}{2} < \gamma_1$; define D by $\gamma_0 < \gamma < \gamma_1$, $|\delta| < \delta_0$; and write

$$\phi_n(c) = \sqrt{\left(\frac{c}{\pi n}\right)} \sum e^{-ch^2/n} A_{n+h}.$$

We may suppose that $|a_n| < n^{-\frac{1}{2}}$ for large n.

Since (as we have just proved) $A_n = O(1)$, $\phi_n(c)$ is, for each n, an analytic function of c regular in D. Next,

$$|\phi_n(c)| = O\left\{\sqrt{\left(\frac{|c|}{n}\right)} \sum e^{-\gamma h^2/n}\right\} = O\left\{\sqrt{\left(\frac{\gamma}{n}\right)} \int e^{-\gamma t^2/n}\, dt\right\} = O(1),$$

uniformly for c in D and all n. By Theorems 151 and 155, $\phi_n(c) \to A$ for every c on the stretch $(\gamma_0, \frac{1}{2})$ of the real axis. It follows from Vitali's theorem that $\phi_n(c) \to A$ for all c of D. In particular, since $\gamma_0 > 0$ and γ_1 are arbitrary, $\sum a_n = A\ (e, c)$ for all positive c.

Thus (9.13.2) is true for all positive c. We may restrict the summation to $|h| < n^{\zeta}$, and then

$$|A_{n+h} - A_n| \leqslant n^{-\frac{1}{2}}|h|$$

for large n. Hence

$$\sqrt{\left(\frac{c}{\pi n}\right)} \left| \sum_{|h|<n^{\zeta}} e^{-ch^2/n}(A_n - A_{n+h}) \right| \leqslant \frac{1}{n}\sqrt{\left(\frac{c}{\pi}\right)} \sum e^{-ch^2/n}|h|$$

$$\leqslant \frac{1}{n}\sqrt{\left(\frac{c}{\pi}\right)} \left\{ \int_{-\infty}^{\infty} e^{-ct^2/n}|t|\, dt + \sqrt{\left(\frac{2n}{ec}\right)} \right\} < \frac{1}{\sqrt{(c\pi)}} + o(1).\dagger$$

It now follows from (9.13.3) that

$$\overline{\lim}\, |A_n - A| \leqslant (c\pi)^{-\frac{1}{2}},$$

and therefore, since c is arbitrary, that $A_n \to A$. This completes the proof.

It is plain that we have also proved

THEOREM 157. *If $\sum a_n$ is summable to A by any of the* (E, q), (e, c), *or* (γ, k) *methods, and $a_n = O(n^{-\frac{1}{2}})$, then $\sum a_n$ converges to A.*

† $e^{-ct^2/n}|t|$ has maxima $\sqrt{(n/2ec)}$ when $t = \pm\sqrt{(n/2c)}$.

9.14. Generalizations. There are generalizations of all these theorems for extensions of Borel's methods such as we encountered in §4.13. There we defined summability (B', α) by

$$\int e^{-t} \sum \frac{a_n t^{n\alpha}}{\Gamma(n\alpha+1)}\, dt = A,$$

and we may define summability (B, α) by

$$e^{-x} \sum A\left(\frac{n}{\alpha}\right) \frac{x^n}{n!} \to A.$$

Thus, when α is an integer k, these assertions are equivalent to the assertions that

$$a_0+0+...+a_1+0+...+a_2+...,$$

where there are $k-1$ zeros between a_n and a_{n+1}, are summable (B') and (B) respectively. We leave it to the reader to prove

THEOREM 158. *If $a_n = o(1)$, then the methods (B, α), (B', α), $(e, \frac{1}{2}\alpha)$ are equivalent.*

THEOREM 159. *Suppose that the parameters c, α, q, k are connected by the relations*

$$c = \tfrac{1}{2}\alpha = \frac{k}{2(1-k)} = \frac{1+q}{2q}$$

(the parameter q being used only when $c > \frac{1}{2}$), and that $a_n = o(1)$. Then the summability of $\sum a_n$ by any one of the methods

$$(e, c), \quad (\mathrm{B}, \alpha), \quad (\mathrm{B}', \alpha), \quad (\mathrm{E}, q), \quad (\gamma, k)$$

implies its summability, to the same sum, by any of the others. Summability for a particular c, α, k, or q implies summability for any smaller positive c, α, or k or any larger q.

It is to be expected, after § 9.12 and the theorems referred to there, that the last clause of Theorem 159, with the restriction $a_n = o(1)$, should be far from expressing the full truth about any one of the methods in question. Thus the implications $(\mathrm{E}, q) \to (\mathrm{E}, q')$ for $q' > q$ and $(\gamma, k) \to (\gamma, k')$ for $0 < k' < k$ are true, after Theorems 118 and 153, without reservation; and it can be proved that $(\mathrm{B}', \alpha) \to (\mathrm{B}', \beta)$ for $0 < \beta < \alpha$ whenever the series

$$\sum a_n \frac{t^{\beta n}}{\Gamma(\beta n+1)}$$

converges for all t.

9.15. The series $\sum z^n$. A good deal of light is thrown on the relations between these methods by their application to the geometric series. We summarize the results shortly.

(1) *The* (E, q) *method.* The series is summable inside the circle $(x+q)^2+y^2 = (q+1)^2$ or

$$r = \sqrt{(1+2q+q^2\cos^2\theta)}-q\cos\theta \quad (-\pi \leqslant \theta \leqslant \pi).$$

See § 8.2.

(2) *The* (B, α) *and* (B$'$, α) *methods.* The success of the methods depends upon the formula (d) of p. 197.† The boundary of the region of summability is the curve $r = \{\sec(\theta/\alpha)\}^\alpha$. The region is bounded if $\alpha > 2$ but extends to infinity if $0 < \alpha \leqslant 2$, and tends to the Mittag-Leffler star when $\alpha \to 0$. When $\alpha = 2$, it is the inside of the parabola $y^2 = 4(1-x)$.

(3) *The* (e, c) *method.* The method will succeed if

$$\sqrt{\left(\frac{c}{\pi n}\right)} \sum \frac{1-z^{n+h}}{1-z} e^{-ch^2/n} \to \frac{1}{1-z}$$

(where $h \geqslant -n$), and the question is easily settled by the aid of the formulae for the linear transformation of the theta-functions. We find that the region of summability is

$$r < e^{\sqrt{(\theta^2+4c^2)}-2c} \quad (-\pi \leqslant \theta \leqslant \pi).$$

(4) *The* (γ, k) *method.* If we are to apply this method to $\sum z^n$, with $|z| > 1$, then we must abandon the restriction imposed on the definition in § 9.11. This invalidates the proof of Theorem 153, which ceases to be generally true. The region of summability is defined by the two inequalities

$$|(1-k)z| < 1, \qquad |kz| < |1-(1-k)z|,$$

and diminishes to the interior of the unit circle when k tends either to 0 or to 1.

It will be found that the relations between these various regions, near $z = 1$, are closest when the parameters are connected as in Theorem 159.

9.16. Valiron's methods. Valiron has defined and used a more comprehensive generalization of Borel's method. The general 'integral-function' definition of § 4.12 was as follows. We say that $A_n \to A$ (J) if

$$\frac{1}{J(x)} \sum c_n A_n x^n = \frac{\sum c_n A_n x^n}{\sum c_n x^n} \to A, \tag{9.16.1}$$

where $J(x)$ is an integral function, not a polynomial, with non-negative coefficients c_n. We now suppose that $c_n = e^{-G(n)}$, where

$$G(n) \to \infty, \qquad G'(n) \to \infty, \qquad G''(n) \to 0$$

† Note on § 8.10.

with considerable regularity. The typical cases are

(9.16.2) $$G(n) = Cn^k \quad (C > 0,\ 1 < k < 2)$$

and $G(n) = n\log n - n$: in the last case the method is practically Borel's. We write $H(n)$ for $G''(n)$, so that

$$H(n) = Ck(k-1)n^{k-2}, \quad n^{-1}$$

in the two typical cases. Then Valiron, generalizing the arguments applied to Borel's method by Hardy and Littlewood, proves that, subject to certain conditions of regularity on $G(n)$, $A_n \to A$(J) is equivalent to

(9.16.3) $$\sqrt{\left\{\frac{H(n)}{2\pi}\right\}} \sum e^{-\frac{1}{2}h^2 H(n)} A_{n+h} \to A$$

whenever

(9.16.4) $$A_n = o[\{H(n)\}^{-\frac{1}{2}}];$$

and that this is true, in particular, whenever $a_n = o(1)$.

We shall express (9.16.3) as $A_n \to A$ (V, H). The main interest of the method lies in the Tauberian theorem associated with it. If $\sum a_n$ is summable (V, H), and

(9.16.5) $$a_n = O[\{H(n)\}^{\frac{1}{2}}],$$

then $\sum a_n$ is convergent. When $G(n) = Cn^k$, then (9.16.5) is

(9.16.6) $$a_n = O(n^{\frac{1}{2}k-1}).$$

We confine ourselves here to proving, as is easy, that (9.16.3), with $H(n) = cn^{k-2}$ $(1 < k < 2)$, and

(9.16.7) $$a_n = o(n^{\frac{1}{2}k-1}),$$

imply the convergence of the series. It will plainly be sufficient to prove that

$$n^{\frac{1}{2}k-1} \sum_{|h|<n^\zeta} e^{-\frac{1}{2}ch^2n^{k-2}}(A_{n+h} - A_n) \to 0,$$

where ζ is now a number between $1-\frac{1}{2}k$ and 1; and this is

$$o(n^{k-2} \sum e^{-\frac{1}{2}ch^2n^{k-2}}|h|) = o(1).$$

NOTES ON CHAPTER IX

§§ 9.1–3. The substance of Theorems 137–9 may be found in many books, particularly in books on the mathematical theory of probability.

§ 9.5. Hadamard's theorem was first proved in *J. de M.* (4), 8 (1892), 101–86 (118), and Fabry's generalization in *AEN* (3), 13 (1896), 367–99. There are comparatively simple proofs in Landau, *Ergebnisse*, 76–86. Ostrowski's theorem was proved in *BS* (1921), 557–65: see also *JLMS*, 1 (1926), 251–63, where fuller references are given. Mordell, *JLMS*, 2 (1937), 146–8, gave a particularly simple

proof of Hadamard's theorem; and Estermann, ibid. 7 (1932), 19–20, proved Ostrowski's by a modification of Mordell's method. The proof here is due to Zygmund, *JLMS*, 6 (1931), 162–3.

See also Dienes, ch. 11.

§§ 9.6–7. Hardy and Littlewood, *PLMS* (2), 11 (1913), 1–16. There is a direct proof of the theorem for Euler summability corresponding to Theorem 143 in Knopp, *MZ*, 18 (1923), 125–56 (136–9).

Hardy and Littlewood state Theorem 147, but give the proof in full only when $2\rho+1$ is an integer. See also Lord, *PLMS* (2), 38 (1935), 241–56.

§ 9.8. Hardy and Littlewood, l.c. *supra*, show that

$$\sum n^{-s}e^{Ain^a} \quad (A > 0)$$

is summable (B) for all s if $\frac{1}{2} < a < 1$, but summable only when convergent if $0 < a \leqslant \frac{1}{2}$. They take $A = 1$, and the argument is only sketched in places.

Another example of an ordinary Dirichlet's series summable (B) only when convergent, though summable (C, k) for some k, all over the plane, is

$$1^{-s}+0+0+...-8^{-s}+0+0+...+27^{-s}+0+....$$

See Hardy, *PLMS* (2), 8 (1909), 277–94 (286–9).

Series of the type (iii) are considered by Hardy, *QJM*, 35 (1904), 22–63. Hardy shows in particular that the convergent series $\sum a_n$ in which

$$a_n = (-1)^m m^{-1} \quad (n = m^2), \qquad a_n = 0 \quad (n \neq m^2)$$

is not absolutely summable.

§ 9.9. The analogue of Theorem 149 for B summability was proved by Hardy (l.c. *supra*, 37–42), and Theorem 149 itself by Knopp (l.c. under § 9.6, 150–1). There are generalizations by Hardy and Littlewood [*RP*, 41 (1916), 36–53 (46–7)] and by Knopp (l.c. 151–2).

§§ 9.10–11. The (e,c) and (γ,k) methods were introduced by Hardy and Littlewood in their paper in the *RP*, and the substance of most of the theorems proved here will be found there or in their later paper in *JLMS*, 18 (1943), 194–200. Knopp considers the relations between Euler summability and summability (e,c), but does not state Theorem 152 explicitly.

§ 9.12. The reference to Hyslop is to *PLMS* (2), 41 (1936), 243–56.

§ 9.13. The method is that of Hardy and Littlewood's paper of 1943. For Vitali's theorem see Littlewood, 117, or Titchmarsh, *Theory of functions*, 168.

Theorem 156 was first proved by Hardy and Littlewood in their paper in the *RP*. The condition (9.13.1) was afterwards generalized further by Valiron [*RP*, 42 (1917), 267–84] and R. Schmidt [*Schriften d. Königsberger gelehrten Ges.* 1 (1925), 205–56]. The most general form, due to Schmidt, is

$$\underline{\lim}(A_n - A_m) \geqslant 0$$

when $m \to \infty$, $n > m$, $m^{-\frac{1}{2}}(n-m) \to 0$: in particular this condition is satisfied if $a_n > -Hn^{-\frac{1}{2}}$. The proof was simplified by Vijayaraghavan, *PLMS* (2), 27 (1927), 316–26.

There is an alternative method of proof by means of Wiener's 'general Tauberian theorems': see Ch. XII and in particular § 12.15.

§ 9.14. The theorem referred to at the end of the section will be found in

Hardy, *JLMS*, 9 (1934), 153–7, and Good, *PCPS*, 38 (1942), 144–65. Hardy proves the result only when $\beta = \frac{1}{2}\alpha$, Good generally. Thus Hardy proves that if

$$a_0+0+a_1+0+a_2+0+... = s \ (\mathrm{B}',\alpha)$$

and

(a) $$a_\alpha(t) = \sum a_n \frac{t^{n\alpha}}{\Gamma(n\alpha+1)}$$

is convergent for all t, then $a_0+a_1+a_2+... = s$ (B',α). The proof actually shows rather more, viz. that if the series (a) converges for small t, and $a_\alpha(t)$ is regular for all positive t, then

$$\int e^{-t} a_\alpha(t)\,dt = s.$$

This is a generalization of the (B',α) method similar to that of the B′ method in § 8.11.

When $q = 1$, $\alpha = 2$. Thus the (B',α) method most closely connected with the ordinary Euler method is $(\mathrm{B}',2)$.

§ 9.15. The main result under (2) is due to Mittag-Leffler, l.c. under § 8.10.

The (e,c) region is determined by Hyslop, l.c. under § 9.12: he has $c = \frac{1}{2}$.

§ 9.16. For all this see Valiron, l.c. under § 9.13. When $k = 2$,

$$G''(n) = H(n) = Ck(k-1) \neq 0$$

and (9.16.6) becomes $a_n = O(1)$. In this case the method fails to sum $1-1+1-...$ (§ 4.15).

X

MULTIPLICATION OF SERIES

10.1. Formal rules for multiplication. It will be convenient, throughout this chapter, to use the same letter for a series and its sum, whether the sum is a sum in the sense of ordinary convergence or not.

If the series $\sum a_m = A$ and $\sum b_n = B$ are absolutely convergent, then the double series $\sum\sum a_m b_n$ is absolutely convergent, and has the sum AB however its terms are arranged. When one at least of A and B is not absolutely convergent, the convergence, or summability, of the double series will depend upon the rule prescribed for its arrangement, different rules giving different definitions of the 'product' of A and B. The most familiar rule is Cauchy's, in which the double series is summed 'diagonally', by associating together the terms in which $m+n$ has a fixed value. We then write

(10.1.1) $$c_p = \sum_{m+n=p} a_m b_n = \sum a_m b_{p-m} = \sum a_{p-n} b_n,\dagger$$

and define the product series C as

(10.1.2) $$C = \sum c_p.$$

The rule is suggested by the formula $\sum a_m x^m \sum b_n x^n = \sum c_p x^p$ for the product of two ascending power series.

Another rule is important in the theory of 'ordinary' Dirichlet series. We suppose that $a_0 = 0$, $b_0 = 0$, and associate together the terms in which mn has a fixed value p, so that

(10.1.3) $$c_p = \sum_{mn=p} a_m b_n = \sum_{d|p} a_d b_{p/d} = \sum_{d|p} a_{p/d} b_d,$$

the summation in the last sums extending over the divisors of p. This rule is suggested by the formula $\sum a_m m^{-s} \sum b_n n^{-s} = \sum c_p p^{-s}$; and there are other modes of multiplication associated with Dirichlet series $\sum a_m e^{-\lambda_m s}$ of more general types.

In this chapter we shall concentrate our attention on Cauchy's rule, which we shall discuss very thoroughly, and dismiss the various methods of 'Dirichlet' multiplication summarily. We shall also say something about the corresponding problems for series infinite in both directions.

10.2. The classical theorems for multiplication by Cauchy's rule. The three classical theorems of the subject, due to Cauchy, Mertens, and Abel respectively, are as follows.

† Here, and often in what follows, we use the convention of § 5.4.

THEOREM 160. *If A and B are absolutely convergent, then C is absolutely convergent, and $C = AB$.*

This is a corollary of the absolute convergence of the double series $\sum\sum a_m b_n$.

THEOREM 161. *If A is absolutely convergent, and B convergent, then C is convergent, and $C = AB$.*

If, as usual, $A_m = a_0+a_1+...+a_m$, and similarly with other letters, then

$$C_p = \sum_{q\leqslant p}\sum_{m+n=q} a_m b_n = \sum_{m+n\leqslant p} a_m b_n = \sum a_m B_{p-m}. \tag{10.2.1}$$

We may write this as $\sum a_m \beta_{m,p}$, where $\beta_{m,p}$ is B_{p-m} if $m \leqslant p$ and 0 if $m > p$. Since $\beta_{m,p}$ is uniformly bounded, and A is absolutely convergent, the series $\sum a_m \beta_{m,p}$ converges uniformly in p; and so

$$\lim_{p\to\infty}\sum a_m \beta_{m,p} = \sum a_m \lim_{p\to\infty}\beta_{m,p} = B\sum a_m = AB.$$

We may add that, if A is not absolutely convergent, then there are convergent B for which C is divergent. For if $C_p = a_0 B_p+...+a_p B_0$ tends to a limit whenever B_n tends to a limit, then $\sum |a_n| < H$, by Theorem 1.

THEOREM 162. *If A, B, and C are all convergent then $C = AB$.*

The power series $a(x) = \sum a_n x^n$,... are absolutely convergent for $0 < x < 1$, and $a(x)b(x) = c(x)$, by Theorem 160. Making $x \to 1$, it follows from Theorem 55 that $AB = C$.

10.3. Multiplication of summable series. There are important generalizations of the preceding theorems for series summable (C, k). The fundamental theorem is Cesàro's Theorem 164: but we begin with

THEOREM 163. *If A, B, and C are all summable* (C, k) *for some k, then $C = AB$.*

In fact the proof of Theorem 162 shows that $C = AB$ whenever all three series are summable (A).

THEOREM 164. *If $r > -1$, $s > -1$, A is summable* (C, r), *and B summable* (C, s), *then C is summable* (C, $r+s+1$), *and $C = AB$.*

We have

$$(1-x)^{-r-s-2}c(x) = (1-x)^{-r-1}a(x).(1-x)^{-s-1}b(x)$$

for $|x| < 1$, and therefore, in the notation of § 5.4,

$$\sum C_p^{r+s+1}x^p = \sum A_m^r x^m \sum B_n^s x^n.$$

Hence, equating coefficients,

$$C_p^{r+s+1} = A_0^r B_p^s + A_1^r B_{p-1}^s + \dots + A_p^r B_0^s. \tag{10.3.1}$$

Since
$$A_m^r \sim A\binom{m+r}{r}, \qquad B_n^s \sim B\binom{n+s}{s},$$
it follows from (10.3.1) and Theorem 41 that
$$C_p^{r+s+1} \sim AB\binom{p+r+s+1}{r+s+1},$$
i.e. that C is summable $(C, r+s+1)$ to AB.

We can naturally prove (10.3.1) without using power series. Since
$$A_k^r = \sum_{m=0}^{k} \binom{k-m+r}{r} a_m, \qquad B_{p-k}^s = \sum_{n=0}^{p-k} \binom{p-k-n+s}{s} b_n,$$
the coefficient of $a_m b_n$ in the right-hand side of (10.3.1) is
$$\sum_{\nu=m}^{p-n} \binom{\nu-m+r}{r}\binom{p-n-\nu+s}{s} = \sum_{k=0}^{p-m-n} \binom{k+r}{r}\binom{p-m-n-k+s}{s}$$
$$= \binom{p-m-n+r+s+1}{r+s+1},$$
by (5.6.10); and this is the corresponding coefficient in C_p^{r+s+1}.

In particular, when $r = s = 0$, (10.3.1) becomes
$$C_p^1 = A_0 B_p + A_1 B_{p-1} + \dots + A_p B_0.$$
It is easy to verify this directly, and to deduce, without appealing to Theorem 41 in its general form, that $A_m \to A$ and $B_n \to B$ imply $C_p^1 \sim ABp$. It follows that C is summable (C, 1), to sum AB, whenever A and B are convergent, and so that, if C also is convergent, its sum is necessarily AB. We thus obtain a simple proof of Theorem 162 independent of the theory of power series.

We add two negative theorems showing that Theorem 164 is a best possible theorem of its kind.

THEOREM 165. *The hypotheses of Theorem 164 do not imply that C is summable (C, k) for any k less than $r+s+1$.*

Take $\rho = r-\delta$, $\sigma = s-\delta$, where δ is positive and small, and
$$a_m = (-1)^m(m+1)^\rho, \qquad b_n = (-1)^n(n+1)^\sigma.$$
Then A and B are summable (C, r) and (C, s), by Theorem 81. But
$$(-1)^p c_p = \sum (m+1)^\rho (p-m+1)^\sigma \sim \frac{\Gamma(\rho+1)\Gamma(\sigma+1)}{\Gamma(\rho+\sigma+2)} p^{\rho+\sigma+1},$$
by Theorem 41, and so, after Theorem 46, C is not summable $(C, \rho+\sigma+1)$, i.e. $(C, r+s+1-2\delta)$.

THEOREM 166. *The result of Theorem 164 is not true for $r = -1$, $s > -1$.*

Take
$$a_m = (-1)^m(m+2)^{-1}\{\log(m+2)\}^{-\alpha}, \qquad b_n = (-1)^n(n+2)^s\{\log(n+2)\}^{-\beta},$$

where $\alpha > 0, \beta > 0, \alpha+\beta < 1$. Then A is summable $(\mathrm{C}, -1)$ and B is summable (C, s). But

$$(-1)^p c_p = \sum_{m=0}^{p} \frac{1}{(m+2)\{\log(m+2)\}^\alpha} \frac{(p-m+2)^s}{\{\log(p-m+2)\}^\beta}$$
$$> \frac{Mp^s}{(\log p)^\beta} \sum_{m=0}^{\frac{1}{2}p} \frac{1}{(m+2)\{\log(m+2)\}^\alpha} > Np^s(\log p)^{1-\alpha-\beta},$$

for constant M, N, and large p. Hence $c_p \neq o(p^s)$, and C is not summable (C, s).

Our next theorem is a generalization of Mertens's Theorem 161, to which it reduces for $r = 0$.

THEOREM 167. *If $r > 0$, A is absolutely convergent, and B is summable (C, r), then C is summable (C, r).*

Here $a(x)\{(1-x)^{-r-1}b(x)\} = (1-x)^{-r-1}c(x)$, and so

$$C_p^r = a_0 B_p^r + a_1 B_{p-1}^r + \ldots + a_p B_0^r.$$

We may write this as

$$\binom{p+r}{r}^{-1} C_p^r = \sum a_m \beta_{m,p},$$

where
$$\beta_{m,p} = \binom{p+r}{r}^{-1} B_{p-m}^r$$

for $m \leqslant p$ and $\beta_{m,p} = 0$ for $m > p$. Since, for $m \leqslant p$,

$$|\beta_{m,p}| \leqslant \binom{p-m+r}{r}^{-1} |B_{p-m}^r|,$$

$\beta_{m,p}$ is uniformly bounded, so that $\sum a_m \beta_{m,p}$ converges uniformly in p. Also $\binom{p-m+r}{r} \sim \binom{p+r}{r}$, for each m, when $p \to \infty$, and so

$$\sum a_m \beta_{m,p} \to \sum a_m \lim \beta_{m,p} = B \sum a_m = AB.$$

THEOREM 168. *The result of Theorem 167 is not true when $r < 0$.*

Take $a_m = (-1)^m \alpha_m$, $b_n = (-1)^n \beta_n$, where α_m and β_n are positive and $\sum \alpha_m < \infty$. Then

$$|c_p| = \sum \alpha_m \beta_{p-m} > \alpha_p \beta_0 \; (p > 0),$$

and $c_p = o(p^r)$ involves $\alpha_p = o(p^r)$, which is not, for any negative r, a consequence of the convergence of $\sum \alpha_p$.

10.4. Another theorem concerning convergence. We now consider a group of theorems of which the simplest is

THEOREM 169. *If A and B are convergent and*

(10.4.1) $$a_m = O(m^{-1}), \qquad b_n = O(n^{-1}),$$

then C is convergent (and $C = AB$).

In fact A and B are both summable $(C, -1+\delta)$, for any positive δ, by Theorem 45; and therefore, by Theorem 164, C is summable $(C, -1+2\delta)$, and *a fortiori* convergent.

THEOREM 170. *If A and B are convergent; $\eta(x)$ and $\zeta(x)$ are positive and tend to infinity with x; $\eta(x)+\zeta(x) = x$; and*

$$\sum_{\eta(x)}^{x} |a_m| = O(1), \qquad \sum_{\zeta(x)}^{x} |b_n| = O(1); \tag{10.4.2}$$

then C is convergent (and $C = AB$).†

We suppose that $0 < y < \eta$, $0 < z < \zeta$, and divide the triangle T of the (m, n)-plane in which $m+n \leqslant x$ into regions T_1, T_2, T_3, T_4, T_5 as shown in Fig. 2. We may suppose x, η, ζ, y, and z non-integral,

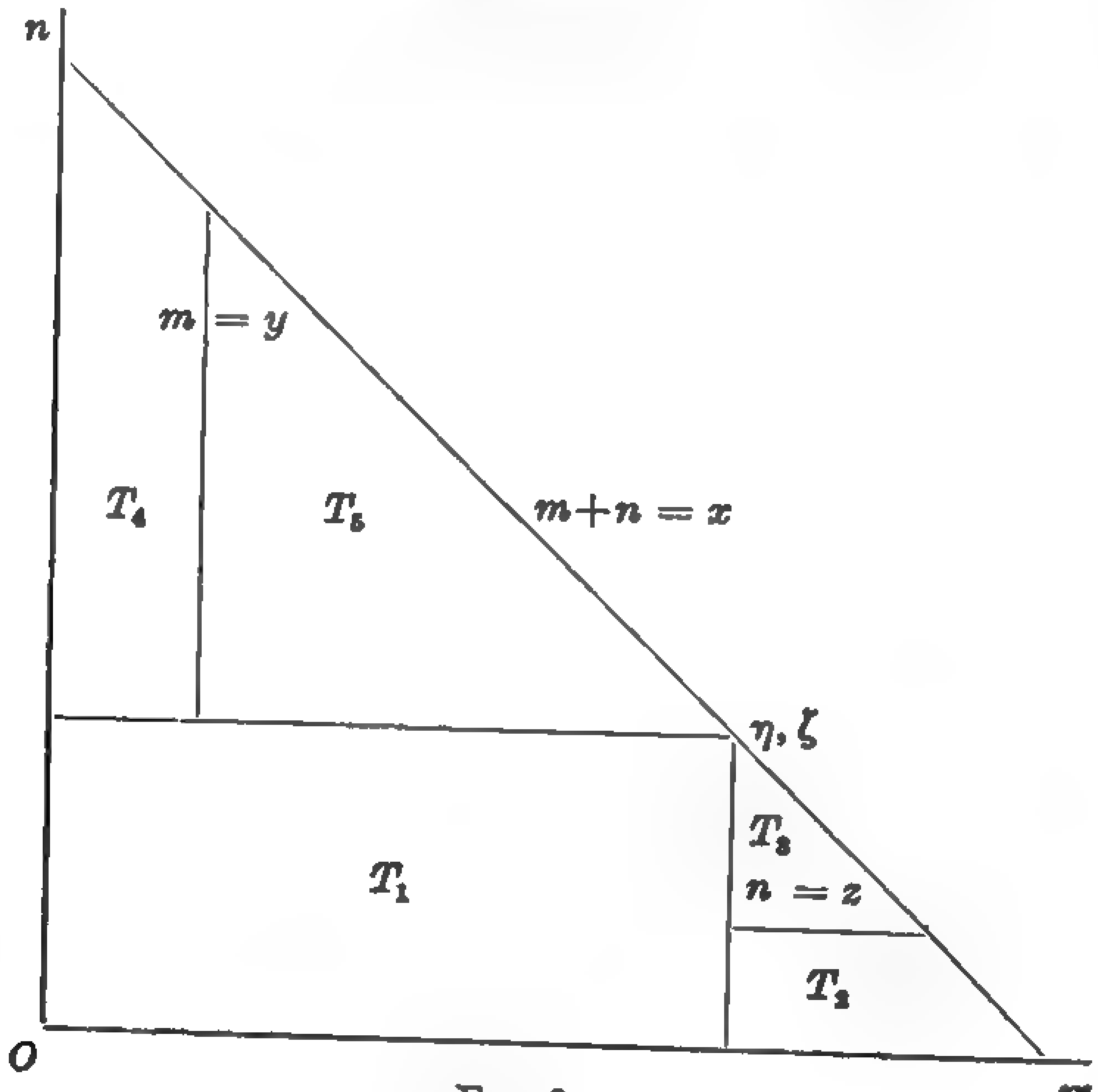

FIG. 2.

so that there are no lattice points on the lines dividing the regions. If $A(x)$,... are the sum-functions of A,..., and Σ_1,... are the sums $\sum\sum a_m b_n$ over T_1,..., then

$$C(x) = \Sigma_1+\Sigma_2+\Sigma_3+\Sigma_4+\Sigma_5.$$

Also
$$\Sigma_1 = \sum_{m<\eta} a_m \sum_{n<\zeta} b_n = A(\eta)B(\zeta) \to AB.$$

† If η or ζ does not tend to infinity, then one or other series is absolutely convergent, and the theorem is included in Theorem 161.

It is therefore sufficient to prove that $\Sigma_2+\Sigma_3 \to 0$ and $\Sigma_4+\Sigma_5 \to 0$, when y and z are chosen appropriately.

We choose z so that

$$\left|\sum_{n=\nu_1}^{\nu_2} b_n\right| < \epsilon \quad (\nu_2 > \nu_1 > z).$$

Then

$$|\Sigma_3| = \left|\sum_{m=\eta}^{x-z} a_m \sum_{n=z}^{x-m} b_n\right| \leqslant \sum_{m=\eta}^{x-z} |a_m| \left|\sum_{n=z}^{x-m} b_n\right| \leqslant \epsilon \sum_{m=\eta}^{x} |a_m| < K\epsilon,$$

for a constant K, by (10.4.2). And

$$|\Sigma_2| = \left|\sum_{n=0}^{z} b_n \sum_{m=\eta}^{x-n} a_m\right| \leqslant \sum_{n=0}^{z} |b_n| \left|\sum_{m=\eta}^{x-n} a_m\right| \to 0$$

when z is fixed and x and η tend to infinity. Hence $\Sigma_2+\Sigma_3 \to 0$; and we can prove that $\Sigma_4+\Sigma_5 \to 0$ similarly.

When the conditions of Theorem 169 are satisfied, we can take $\eta = \zeta = \frac{1}{2}x$, since

$$\sum_{\frac{1}{2}x \leqslant m \leqslant x} |a_m| = O\left(\sum_{\frac{1}{2}x \leqslant m \leqslant x} \frac{1}{m}\right) = O(1).$$

There are other interesting special cases.

THEOREM 171. *If A and B are convergent,*

(10.4.3) $$a_m = O(m^{-\delta}) \quad (\delta > 0), \qquad b_n = O\left(\frac{1}{n \log n}\right),$$

then C is convergent.

We may plainly suppose $\delta < 1$. We take $\eta = x - x^\delta$, $\zeta = x^\delta$. Then

$$\sum_{\eta}^{x} |a_m| = O(x^\delta . x^{-\delta}) = O(1),$$

$$\sum_{\zeta}^{x} |b_n| = O\left(\sum_{x^\delta}^{x} \frac{1}{n \log n}\right) = O(\log\log x - \log\log x^\delta) = O(1).$$

10.5. Further applications of Theorem 170. (1) Theorem 170 enables us to prove a 'one-sided' extension of Theorem 169.

THEOREM 172. *If a_m and b_n are real, A and B are convergent, and $a_m > -Km^{-1}$, $b_n > -Ln^{-1}$, with constant K and L, then C is convergent.*

If a_m^+ and a_m^- are the positive and negative a_m, and

$$A(x) = \sum_{m \leqslant x} a_m, \qquad A^+(x) = \sum_{m \leqslant x} a_m^+, \qquad A^-(x) = \sum_{m \leqslant x} a_m^-,$$

then
$$A(x) = A^+(x) + A^-(x),$$

$$\sum_{m \leqslant x} |a_m| = A^+(x) - A^-(x) = A(x) - 2A^-(x),$$

$$\sum_{\frac{1}{2}x < m \leqslant x} |a_m| = A(x) - A(\tfrac{1}{2}x) - 2\{A^-(x) - A^-(\tfrac{1}{2}x)\}.$$

But $A(x) = O(1)$ and

$$A^-(x) - A^-(\tfrac{1}{2}x) = \sum_{\frac{1}{2}x<m\leqslant x} a_m^- > -K \sum_{\frac{1}{2}x<m\leqslant x} m^{-1},$$

which is bounded, so that a_m satisfies (10.4.2). The same argument applies to b_n, and the theorem follows.

(2) Another application is to the proof of

THEOREM 173. *If A and B are convergent, $p > 1$, $q > 1$, and $\sum m^{p-1}|a_m|^p < \infty$, $\sum n^{q-1}|b_n|^q < \infty$, then C is convergent.*

For if $p' = p/(p-1)$, and all the summations are over $(\frac{1}{2}x, x)$, then, by Hölder's inequality,

$$\sum |a_m| = \sum m^{1/p'}|a_m| . m^{-1/p'} \leqslant (\sum m^{p-1}|a_m|^p)^{1/p} (\sum m^{-1})^{1/p'} = O(1);$$

and similarly $\sum |b_n| = O(1)$.

10.6. Alternating series. In this section we prove a more elementary theorem concerning series of the familiar alternating type.

THEOREM 174. *If $A = \sum (-1)^m \alpha_m$, $B = \sum (-1)^n \beta_n$, where α_m and β_n are positive and decrease steadily to 0, then, in order that C should be convergent,*

(i) *it is necessary and sufficient that*

(10.6.1) $$\gamma_p = \alpha_0\beta_p + \alpha_1\beta_{p-1} + \ldots + \alpha_p\beta_0 \to 0;$$

(ii) *it is necessary and sufficient that*

(10.6.2) $$(\alpha_0+\alpha_1+\ldots+\alpha_p)\beta_p \to 0, \qquad (\beta_0+\beta_1+\ldots+\beta_p)\alpha_p \to 0;$$

(iii) *it is sufficient that* $\sum \alpha_p\beta_p < \infty$;

(iv) *it is necessary that* $\sum (\alpha_p\beta_p)^{1+\delta} < \infty$ *for every positive* δ.

(i) If we write

$$A = A_m + (-1)^{m+1}\rho_m, \qquad B = B_n + (-1)^{n+1}\sigma_n,$$

then $0 \leqslant \rho_m \leqslant \alpha_m$, $0 \leqslant \sigma_n \leqslant \beta_n$. Also

$$C_p = \sum a_m B_{p-m} = BA_p + (-1)^p \sum \alpha_m \sigma_{p-m} = BA_p + R_p,$$

where $|R_p| \leqslant \sum \alpha_m \beta_{p-m} = \gamma_p$. Hence (10.6.1) is sufficient, and it is obvious that it is necessary.

(ii) Next,

$$\gamma_p \geqslant (\alpha_0+\ldots+\alpha_p)\beta_p, \qquad \gamma_p \geqslant (\beta_0+\ldots+\beta_p)\alpha_p,$$

so that the conditions (10.6.2) are necessary; and

$$\gamma_p \leqslant (\alpha_0+\ldots+\alpha_q)\beta_q + (\beta_0+\ldots+\beta_q)\alpha_q,$$

where $q = [\frac{1}{2}p]$, so that they are also sufficient.

(iii) If $0 < q < p$ then

$$\alpha_p(\beta_0+\beta_1+\ldots+\beta_p) \leqslant \alpha_p(\beta_0+\ldots+\beta_q) + \alpha_{q+1}\beta_{q+1} + \ldots + \alpha_p\beta_p = S_1 + S_2,$$

say. If $\sum \alpha_p\beta_p < \infty$, we can choose q so that $S_2 < \epsilon$ for all $p > q$; and

$S_1 \to 0$ when q is fixed and $p \to \infty$. Hence $\alpha_p(\beta_0+...+\beta_p) \to 0$, and similarly $\beta_p(\alpha_0+...+\alpha_p) \to 0$. Thus the condition is sufficient. That it is not necessary is shown by the example in which $\alpha_n = \beta_n = (n \log n)^{-\frac{1}{2}}$ for $n \geqslant 2$: in this case $\gamma_n \to 0$, so that C is convergent.

(iv) Finally, $\quad (\alpha_0+\alpha_1+...+\alpha_p)\beta_p \geqslant p\alpha_p \beta_p.$

If C is convergent, then the left-hand side tends to 0, by (ii), and so $\alpha_p \beta_p = o(p^{-1})$. Hence $\sum (\alpha_p \beta_p)^{1+\delta} < \infty$. The example

$$\alpha_n = \beta_n = (n+1)^{-\frac{1}{2}}$$

shows that the condition is not sufficient.

10.7. Formal multiplication. We have assumed in all of the preceding theorems that both A and B are convergent or summable. There is another type of theorem, particularly important in the theory of trigonometrical series, in which one series, say A, is almost arbitrary, while the other is severely restricted. The conclusion is then that C_p behaves 'very much like' BA_p.

If

$$C_p - BA_p \to 0 \tag{10.7.1}$$

when $p \to \infty$, then we shall say that C is *equi-convergent* with $B(A)$. In this case, if A is convergent or summable, then so is C (with sum AB).

THEOREM 175. *If* $a_m = o(1)$, $\sum n|b_n| < \infty$, *then C is equi-convergent with* $B(A)$.

We use as a lemma

THEOREM 176. *If* $a_m = o(1)$, $\sum |\rho_n| < \infty$, *then*

$$\sigma_p = a_0 \rho_p + a_1 \rho_{p-1} + ... + a_p \rho_0 = o(1).$$

This is trivial; for we can choose P so that

$$|a_m| < \epsilon \quad (m \geqslant \tfrac{1}{2}P), \qquad \sum_{n \geqslant \frac{1}{2}P} |\rho_n| < \epsilon,$$

$$|\sigma_p| \leqslant \operatorname*{Max}_{m \leqslant \frac{1}{2}p} |a_m| \sum_{\frac{1}{2}p}^{p} |\rho_n| + \operatorname*{Max}_{\frac{1}{2}p < m \leqslant p} |a_m| \sum_{0}^{\frac{1}{2}p} |\rho_n| \leqslant \epsilon(\operatorname{Max}|a_m| + \sum |\rho_n|)$$

for $p \geqslant P$.

To deduce Theorem 175, we write $B = B_n + \beta_n$, so that

$$C_p = a_0 B_p + a_1 B_{p-1} + ... + a_p B_0 = BA_p - a_0 \beta_p - ... - a_p \beta_0,$$

where

$$\beta_p = \sum_{n>p} b_n, \qquad |\beta_p| \leqslant \sum_{n>p} |b_n|, \qquad \sum |\beta_p| \leqslant \sum_p \sum_{n>p} |b_n| = \sum n|b_n|.$$

Then, after Theorem 176,

$$a_0\beta_p+a_1\beta_{p-1}+\dots+a_p\beta_0 \to 0, \qquad C_p-BA_p \to 0.$$

It is plain that if a_m and b_n are functions of a variable x, $a_m \to 0$ uniformly, and $\sum n|b_n|$ is uniformly convergent, then $C_p-BA_p \to 0$ uniformly.

There are generalizations of Theorem 175 for (C, k) summability, a little more complex in form. We leave it to the reader to prove

THEOREM 177. *If $a_m = o(m)$, $\sum n^2|b_n| < \infty$, then*

$$C_p^1-BA_p^1+B^*A_p = o(p),$$

where $B^* = \sum nb_n$. If also $A_m = o(m)$, and in particular if $a_m = o(1)$, then $C_p^1-BA_p^1 = o(p)$, i.e. C and $B(A)$ are equi-summable $(C, 1)$.

10.8. Multiplication of integrals. We now state the theorems for integrals which correspond to the more important of Theorems 160–70. The proofs follow the same lines and we do not give details, noting only the points where there are material differences. These arise from the absence of any 'limitation theorem' corresponding to Theorem 46. In particular, the convergence of $A = \int a(x)\,dx$ does not imply the convergence of $\int e^{-\delta x}|a(x)|\,dx$ for positive δ.

We define C by $C = \int c(x)\,dx$, where

$$c(x) = \int a(t)b(x-t)\,dt = \int a(x-t)b(t)\,dt \tag{10.8.1}$$

(with the convention of § 5.6 concerning limits).

THEOREM 178. *If A and B are absolutely convergent then C is absolutely convergent, and $C = AB$.*

THEOREM 179. *If A is absolutely convergent and B is convergent, then C is convergent, and $C = AB$.*

THEOREM 180. *If all three integrals are convergent, then $C = AB$.*

THEOREM 181. *If $r > -1$, $s > -1$, A is summable (C, r), and B summable (C, s), then C is summable $(C, r+s+1)$, and $C = AB$.*

THEOREM 182. *If $r > 0$, A is absolutely convergent, and B is summable (C, r), then C is summable (C, r), and $C = AB$.*

THEOREM 183. *If A and B are convergent, η and ζ satisfy the conditions of Theorem 170, and*

$$\int_\eta^x |a(t)|\,dt = O(1), \qquad \int_\zeta^x |b(t)|\,dt = O(1), \tag{10.8.2}$$

then C is convergent, and $C = AB$.

We need only add the following remarks.

(i) We must deduce Theorem 180 from Theorem 181 with $r = s = 0$, and not attempt to imitate the argument of § 10.2.

(ii) The proof of Theorem 181 depends on the identity

(10.8.3)
$$C_{r+s+1}(x) = \int A_r(t)B_s(x-t)\,dt,$$

where, for example,

$$A_r(x) = \frac{1}{\Gamma(r+1)}\int (x-t)^r a(t)\,dt;$$

and that of Theorem 182 on

(10.8.4)
$$C_r(x) = \int a(t)B_r(x-t)\,dt.$$

To prove (10.8.3) we observe that

$$\int_0^x A_r(t)B_s(x-t)\,dt = \frac{1}{\Gamma(r+1)\Gamma(s+1)}\int_0^x dt\int_0^t a(u)(t-u)^r\,du\int_0^{x-t} b(v)(x-t-v)^s\,dv.$$

Here $0 \leqslant u \leqslant t \leqslant x$, $0 \leqslant v \leqslant x-t \leqslant x$. When u and v are fixed, t runs from u to $x-v$: when u is fixed, v runs from 0 to $x-u$. Hence the triple integral is

$$\frac{1}{\Gamma(r+1)\Gamma(s+1)}\int_0^x a(u)\,du\int_0^{x-u} b(v)\,dv\int_u^{x-v} (t-u)^r(x-t-v)^s\,dt$$

$$= \frac{1}{\Gamma(r+s+2)}\int_0^x a(u)\,du\int_0^{x-u} (x-u-v)^{r+s+1}b(v)\,dv$$

$$= \frac{1}{\Gamma(r+s+2)}\int_0^x a(u)\,du\int_u^x (x-w)^{r+s+1}b(w-u)\,dw$$

$$= \frac{1}{\Gamma(r+s+2)}\int_0^x (x-w)^{r+s+1}\,dw\int_0^w a(u)b(w-u)\,du = C_{r+s+1}(x).$$

When $r \geqslant 0$, $s \geqslant 0$, the argument is valid, by Fubini's theorem, for all integrable $a(x)$ and $b(x)$. If $a(x)$ and $b(x)$ are bounded in every finite interval $(0, X)$, it is valid for $r > -1$, $s > -1$: in other cases some reservations are needed.†

The proof of (10.8.4) is similar but a little simpler.

10.9. Euler summability. We must now consider the problem of multiplication for series summable by Euler's and Borel's methods. We recall the definition of summability (E, q): if, for small x and y,

$$f(x) = \sum a_n x^{n+1} = \sum a_n\left(\frac{y}{1-qy}\right)^{n+1} = \sum a_m^{(q)}\{(q+1)y\}^{m+1},$$

and $\sum a_n^{(q)} = A$, then $\sum a_n$ is summable (E, q) to sum A. We have to add a further definition: if $\sum a_n^{(q)}$ is summable (C, k), then we say that $\sum a_n$ is summable $(\mathrm{E}, q;\ \mathrm{C}, k)$.

† Compare the note on §§ 5.14–15.

THEOREM 184. *If* $\sum a_m = A$ (E, q), $\sum b_n = B$ (E, q), *then*

$$\sum c_p = AB \ \ (\mathrm{E}, q;\ \mathrm{C}, 1).$$

If

$$f(x) = \sum a_m^{(q)}(q+1)^{m+1}y^{m+1}, \qquad g(x) = \sum b_n^{(q)}(q+1)^{n+1}y^{n+1},$$
$$h(x) = \sum c_p^{(q)}(q+1)^{p+1}y^{p+1},$$

then the series are absolutely convergent for small x and y, and $f(x)g(x) = xh(x)$. Hence

$$\sum c_p^{(q)}(q+1)^p y^p = (q+1)(1-qy)\sum a_m^{(q)}(q+1)^m y^m \sum b_n^{(q)}(q+1)^n y^n,$$

(10.9.1) $$\sum c_p^{(q)}Y^p = (q+1)\sum a_m^{(q)}Y^m \sum b_n^{(q)}Y^n - qY\sum a_m^{(q)}Y^m \sum b_n^{(q)}Y^n,$$

where $Y = (q+1)y$. It follows that

$$c_p^{(q)} = (q+1)\sum_{m+n=p} a_m^{(q)}b_n^{(q)} - q\sum_{m+n=p-1} a_m^{(q)}b_n^{(q)} = (q+1)G_p - qG_{p-1},$$

say (with $G_{-1} = 0$). Since $\sum a_m^{(p)} = A$, $\sum b_n^{(q)} = B$, it follows from Theorem 164 that $\sum G_p = AB$ (C, 1), and from Theorem 47 that

$$\sum G_{p-1} = 0 + G_0 + \dots = AB \ \ (\mathrm{C}, 1).$$

Hence $$\sum c_p^{(q)} = (q+1-q)AB = AB \ \ (\mathrm{C}, 1).$$

It is plain that, if we use Mertens's theorem instead of Cesàro's, we obtain

THEOREM 185. *If* $\sum a_m = A$ (|E, q|), *i.e. if* $\sum a_m^{(q)}$ *converges absolutely to* A, *and* $\sum b_n = B$ (E, q), *then* $\sum c_p = AB$ (E, q).

On the other hand, if we suppose A, B, C all summable (E, q), and make $Y \to 1$ in (10.9.1), we obtain $C = \sum c_p^{(q)} = AB$, and so

THEOREM 186. *If* A, B, C *are all summable* (E, q), *then* $C = AB$.

10.10. Borel summability. The facts concerning Borel summability are a little more complex, since there are two definitions, and since the series $a_0+a_1+\dots$ and $a_1+a_2+\dots$ need not behave similarly. We state our results in terms of the integral definition, leaving the variants for the exponential definition to the reader.

If Borel's integral is summable (C, k), in the sense of § 5.14, we say that $\sum a_n$ is summable (B′; C, k). We begin by proving that

(10.10.1) $$a_0+a_1+a_2+\dots = A \ \ (\mathrm{B}';\ \mathrm{C}, k)$$

implies

(10.10.2) $$a_1+a_2+a_3+\dots = A - a_0 \ \ (\mathrm{B}';\ \mathrm{C}, k+1).$$

In fact, if $a(t) = \sum a_n \frac{t^n}{n!}$, then

(10.10.3)

$$\int^t \left(1-\frac{u}{t}\right)^{k+1} e^{-u}\left(a_1+a_2\frac{u}{1!}+\ldots\right) du = \int_0 \left(1-\frac{u}{t}\right)^{k+1} e^{-u}a'(u)\, du$$

$$= -a_0+\int_0^t \left(1-\frac{u}{t}\right)^{k+1} e^{-u}a(u)\, du+\frac{k+1}{t}\int_0^t \left(1-\frac{u}{t}\right)^{k} e^{-u}a(u)\, du.$$

Also (10.10.1) implies

$$a_0+a_1+a_2+\ldots = A \quad (\mathrm{B}';\ \mathrm{C}, k+1).$$

Hence the second term on the right of (10.10.3) tends to A, while the last term is $O(t^{-1})$; and (10.10.2) follows.

We can now prove

THEOREM 187. *If $\sum a_m$ and $\sum b_n$ are summable* (B′) *to A and B, then*

(10.10.4) $$0+c_0+c_1+c_2+\ldots = AB \quad (\mathrm{B}';\ \mathrm{C}, 1)$$

and

(10.10.5) $$c_0+c_1+c_2+\ldots = AB \quad (\mathrm{B}';\ \mathrm{C}, 2).$$

By Theorem 181, with $r = s = 0$,

$$\int_0^\infty e^{-x}\, dx \int_0^x a(t)b(x-t)\, dt = \int_0^\infty dx \left\{\int_0^x e^{-t}a(t)\,.\,e^{-x+t}b(x-t)\, dt\right\} = AB\ (\mathrm{C}, 1).$$

The inner integral on the left is

$$\sum_m \sum_n \frac{a_m b_n}{m!\,n!}\int_0^x t^m(x-t)^n\, dt = \sum_m \sum_n \frac{a_m b_n}{(m+n+1)!}x^{m+n+1} = \sum_p c_p \frac{x^{p+1}}{(p+1)!}.$$

Hence $$\int e^{-x}\sum c_p \frac{x^{p+1}}{(p+1)!}\, dx = AB \quad (\mathrm{C}, 1),$$

which is (10.10.4); and (10.10.5) is now a corollary.

It is plain that, using Theorems 179 and 180, respectively, instead of Theorem 181, we can prove the following two theorems.

THEOREM 188. *If (in addition to the hypotheses of Theorem 187) $\int e^{-t}|a(t)|\, dt < \infty$ (i.e. if A is absolutely summable), then*

$$0+c_0+c_1+\ldots = AB \ (\mathrm{B}'), \qquad c_0+c_1+c_2+\ldots = AB \ (\mathrm{B}';\ \mathrm{C}, 1).$$

THEOREM 189. *If $\sum a_m$, $\sum b_n$, and $\sum c_p$ are all summable* (B′), *then $C = AB$.*

10.11. Dirichlet multiplication. Suppose that $\lambda_0 \geqslant 0$, $\mu_0 \geqslant 0$, and that the sequences (λ_m) and (μ_n) increase strictly to ∞; that (ν_p) is the sequence $(\lambda_m+\mu_n)$ arranged in ascending order, equal sums $\lambda_m+\mu_n$ being regarded as giving one ν_p; and that

$$c_p = \sum_{\lambda_m+\mu_n=\nu_p} a_m b_n.$$

Then we call $C = \sum c_p$ the general Dirichlet product of A and B. If $\lambda_m = m$, $\mu_n = n$, then the rule reduces to Cauchy's; if $\lambda_m = \log m$, $\mu_n = \log n$, to the rule defined by (10.1.3), which has many applications in the theory of numbers. The general theory of Dirichlet multiplication demands a detailed study of Riesz's 'typical means' defined in § 4.16. We confine ourselves here to the generalization of Mertens's theorem.

THEOREM 190. *If A is absolutely convergent and B convergent, then the general Dirichlet product C of A and B is convergent, and $C = AB$.*

For

$$C_P = \sum_{p\leqslant P} c_p = \sum_{\lambda_m+\mu_n\leqslant\nu_P} a_m b_n = \sum_{\lambda_m\leqslant\nu_P-\mu_0} a_m \sum_{\mu_n\leqslant\nu_P-\lambda_m} b_n = \sum_{\lambda_m\leqslant\nu_P-\mu_0} a_m B_N,$$

where $N = N(m,p)$ is the largest n for which $\mu_n \leqslant \nu_P-\lambda_m$; i.e.

$$C_P = \sum a_m \beta_{m,P}, \tag{10.11.1}$$

where $$\beta_{m,P} = \sum_{\mu_n\leqslant\nu_P-\lambda_m} b_m \quad (\lambda_m \leqslant \nu_P), \qquad 0 \quad (\lambda_m > \nu_P).$$

Then $\beta_{m,P}$ is uniformly bounded, (10.11.1) uniformly convergent, and

$$C_P \to \sum a_m \lim \beta_{m,P} = B\sum a_m = AB.$$

10.12. Series infinite in both directions. We end this chapter with a short discussion of the multiplication problem for series infinite in both directions. The problem is a good deal more difficult than the problem for ordinary series, since the general term of the product series is usually itself an infinite series. We shall have to consider two different definitions of the sum of a series over $(-\infty,\infty)$, and two different rules for multiplication. We shall find it convenient to vary our conventions concerning sums written without limits. A sum $\sum a_n$ without limits will run over all integral n, and we shall write $\sum^+ a_n$ and $\sum^- a_n$ for sums over positive and negative n, so that

$$\sum a_n = a_0 + \sum{}^+ a_n + \sum{}^- a_n$$

when the series are convergent.

If

$$A_{N,N'} = \sum_{-N'\leqslant n\leqslant N} a_n \to A \tag{10.12.1}$$

when N and N' tend to ∞ independently, then we shall say that A is *unrestrictedly* convergent. In this case the series $\sum^{+} a_n$ and $\sum^{-} a_n$ converge separately, and $a_n \to 0$ when $|n| \to \infty$.

If

(10.12.2) $$A_N = A_{N,N} \to A$$

when $N \to \infty$, we shall say that A is *restrictedly* convergent. In this case $\sum^{+} (a_n + a_{-n})$ is convergent, and $a_n + a_{-n} \to 0$; but there is no limitation on the order of a_n and a_{-n} separately.

We define the product of two series A and B in one or other of two ways.

(1) *Laurent multiplication.* The formal product of two Laurent series

$$A(x) = \sum a_m x^m, \qquad B(x) = \sum b_n x^n,$$

arranged in powers of x, is

$$C(x) = \sum c_p x^p,$$

where

(10.12.3) $$c_p = \sum_{m+n=p} a_m b_n = \sum a_m b_{p-m} = \sum a_{p-n} b_n.\dagger$$

The rule for Laurent multiplication of A and B is obtained by putting $x = 1$, so that C is $\sum c_p$.

(2) *Fourier multiplication.* There is another rule which is particularly adapted for restrictedly convergent series (and which we shall use only for such series). If we write

(10.12.4) $$\begin{cases} A(\theta) = \sum a_m \cos m\theta = \tfrac{1}{2}\alpha_0 + \sum^{+} \alpha_m \cos m\theta, \\ B(\theta) = \sum b_n \cos n\theta \ = \tfrac{1}{2}\beta_0 + \sum^{+} \beta_n \cos n\theta, \end{cases}$$

where

(10.12.5) $$\alpha_m = a_m + a_{-m}, \qquad \beta_n = b_n + b_{-n}$$

(so that $\alpha_0 = 2a_0$, $\beta_0 = 2b_0$, and α_m and β_n are even), multiply $A(\theta)$ and $B(\theta)$ formally, and use the addition formulae for cosines, we obtain

(10.12.6) $$C(\theta) = \sum c_p \cos p\theta = \tfrac{1}{2}\gamma_0 + \sum^{+} \gamma_p \cos p\theta,$$

where $$c_p = \tfrac{1}{2} \sum_{m \pm n = p} a_m b_n,\ddagger$$

(10.12.7)

$$\gamma_0 = 2c_0 = \sum_{m \pm n = 0} a_m b_n, \qquad \gamma_p = c_p + c_{-p} = \tfrac{1}{2} \sum_{m \pm n \pm p = 0} a_m b_n \quad (p > 0).$$

Thus γ_0 is the sum of the products $a_m b_n$ on the two lines $m \pm n = 0$,

† These are sums over $(-\infty, \infty)$: we do not use the convention of §.5.4 here.

‡ It would be equally natural to define c_p with $m \pm n = -p$, but we shall always associate c_p with c_{-p}.

and γ_p, for $p > 0$, is half the sum of the products on the four lines $m \pm n \pm p = 0$, the points of intersection of these lines being in either case counted twice.

The rule for Fourier multiplication is obtained by putting $\theta = 0$ in $A(\theta)$, $B(\theta)$, and $C(\theta)$. Thus

(10.12.8) $$A = \tfrac{1}{2}\alpha_0 + \textstyle\sum^+ \alpha_m, \quad B = \tfrac{1}{2}\beta_0 + \textstyle\sum^+ \beta_n, \quad C = \tfrac{1}{2}\gamma_0 + \textstyle\sum^+ \gamma_p,$$

where

(10.12.9)

$$\gamma_p = \tfrac{1}{2}\alpha_0\beta_p + \tfrac{1}{2}\textstyle\sum^+ \alpha_m(\beta_{m-p} + \beta_{m+p}) = \tfrac{1}{2}\beta_0\alpha_p + \tfrac{1}{2}\textstyle\sum^+ \beta_n(\alpha_{n-p} + \alpha_{n+p}),$$

and in particular

(10.12.10) $$\gamma_0 = \tfrac{1}{2}\alpha_0\beta_0 + \textstyle\sum^+ \alpha_m\beta_m.$$

Our work so far is formal. Whichever definition we adopt, c_p, or γ_p, is defined by infinite series which need not converge. For example, if

$$a_0 = b_0 = 0, \qquad a_m = b_m = (-1)^m|m|^{-\frac{1}{2}} \quad (m \neq 0),$$

then the definition (10.12.3) gives

$$c_p = (-1)^p \textstyle\sum' |m|^{-\frac{1}{2}}|p-m|^{-\frac{1}{2}}$$

(where the dash implies omission of the terms $m = 0$ and $m = p$), and the series diverges for every p.

10.13. The analogues of Cauchy's and Mertens's theorems. It is plain that Cauchy's theorem for two absolutely convergent series stands unchanged, for either rule of multiplication, and we need only consider the analogue of Mertens's theorem. Here we must distinguish between the two rules.

THEOREM 191. *If A is absolutely, and B unrestrictedly, convergent, then the Laurent product C is unrestrictedly convergent, and $C = AB$.*

It is plain first, since A is absolutely convergent and b_n bounded, that the series for every c_p is (absolutely) convergent. Also

$$C_{P,P'} = \sum_{p=-P'}^{P} c_p = \sum_{p=-P'}^{P} \sum_{m+n=p} a_m b_n$$

$$= \sum_{m=-\infty}^{\infty} a_m \sum_{n=-P'-m}^{P-m} b_n = \sum a_m \beta_{m,P,P'},$$

say. Since $\beta_{m,P,P'}$ is uniformly bounded, this series is uniformly convergent, and

$$C_{P,P'} \to \sum a_m \lim \beta_{m,P,P'} = B \sum a_m = AB.$$

The theorem becomes false if 'unrestrictedly' is replaced by 'restrictedly' in hypothesis and conclusion: the hypotheses do not then involve even the existence of c_p. Thus, if

$$a_m = 0\ (m \leqslant 0),\quad a_m = m^{-2}\ (m > 0),\quad b_0 = 0,\quad b_n = -b_{-n} = 2^n\ (n > 0),$$

then A converges absolutely to $\frac{1}{6}\pi^2$ and B converges restrictedly to 0, but

$$c_0 = -\sum{}^{+} n^{-2}2^n = -\infty.$$

The corresponding theorem for restrictedly convergent series is

THEOREM 192. *If A is absolutely convergent and B restrictedly convergent, then the Fourier product C of A and B is convergent, and $C = AB$.*

In this case

$$C_P = \sum_{-P}^{P} c_p = \tfrac{1}{2}\gamma_0 + \sum_{1}^{P} \gamma_p = \tfrac{1}{2}\sum_{D}\sum a_m b_n, \tag{10.13.1}$$

where D is the infinite cross defined by $|m \pm n| \leqslant P$, and products corresponding to points of the square $|m|+|n| \leqslant P$ are counted twice. Now, if

$$u_m = \tfrac{1}{2}\alpha_m,\qquad v_n = \tfrac{1}{2}\beta_n,\qquad U = \sum u_m,\qquad V = \sum v_n,$$

then U is absolutely and V unrestrictedly convergent, so that, by Theorem 191, their Laurent product $W = \sum w_p$ is (unrestrictedly) convergent, and $W = UV$. Also

$$W_P = \sum_{-P}^{P} w_p = \sum_{|m+n| \leqslant P}\sum u_m v_n = \tfrac{1}{4}\sum_{|m+n| \leqslant P}\sum (a_m + a_{-m})(b_n + b_{-n}), \tag{10.13.2}$$

and a moment's consideration shows that $C_P = W_P$. Hence

$$C_P \to UV = AB.$$

10.14. Further theorems. There is a theorem for Laurent multiplication corresponding to Theorem 175, viz.

THEOREM 193. *If $a_n = o(1)$ when $|n| \to \infty$, and $\sum |nb_n| < \infty$, then $C_{P,P} - BA_{P,P'} \to 0$. In particular, if A is unrestrictedly (restrictedly) convergent, then C converges unrestrictedly (restrictedly) to AB,*

and a similar theorem for Fourier multiplication of restrictedly convergent series, which we leave to the reader.

The analogue of Theorem 169 requires more careful consideration. The most satisfactory statement is in terms of Fourier multiplication.

THEOREM 194. *If $\alpha_m = O(|m|^{-1})$, $\beta_n = O(|n|^{-1})$, and A and B are restrictedly convergent, then the Fourier product of A and B converges to AB.*

First, after (10.12.9), γ_p is the sum of two series of the type

$$\sum O\left(\frac{1}{|m|+1}\,\frac{1}{|m-p|+1}\right).$$

These are absolutely convergent, and their sums tend to 0 when $p\to\infty$.

Next, after (10.13.1) and (10.13.2),

$$C_P = \tfrac{1}{4}\sum_{|m+n|\leqslant P}\sum \alpha_m\beta_n = \tfrac{1}{4}(T_1+T_2+T_3),$$

where T_1, T_2, T_3 extend over the ranges D_1, D_2, D_3 defined by

$|m|+|n|\leqslant P;\quad |m+n|\leqslant P,\ m-n>P;\quad |m+n|\leqslant P,\ m-n<-P$

respectively. Now $$\tfrac{1}{4}T_1 = \sum_{D_1^+}\sum \alpha_m\beta_n,$$

where D_1^+ is the positive quarter of D_1; terms on the axes are multiplied by $\frac{1}{2}$, and $\alpha_0\beta_0$ by $\frac{1}{4}$. This sum is the partial sum of the Cauchy product of $\frac{1}{2}\alpha_0+\sum^+\alpha_m$ and $\frac{1}{2}\beta_0+\sum^+\beta_n$, and hence, by Theorem 169, $\frac{1}{4}T_1\to AB$. It is therefore sufficient to prove that T_2 and T_3 tend to 0.

We take T_2. We have

$$T_2 = \sum_{m=1}^{P}\alpha_m\sum_{n=-P-m}^{m-P-1}\beta_n+\sum_{m=P+1}^{\infty}\alpha_m\sum_{n=-P-m}^{P-m}\beta_n$$
$$= \sum_{m=1}^{P}\alpha_m\sum_{n=P-m+1}^{P+m}\beta_n+\sum_{m=P+1}^{\infty}\alpha_m\sum_{n=m-P}^{m+P}\beta_n = V_1+V_2,$$

say. First

$$V_1 = \left(\sum_{m=1}^{\delta P}+\sum_{\delta P}^{P-\eta}+\sum_{P-\eta}^{P}\right)\alpha_m\sum_{n=P-m+1}^{P+m}\beta_n = V_1^{(1)}+V_1^{(2)}+V_1^{(3)}.$$

Here $0<\delta<\frac{1}{2}$ and η is large enough to make

$$\left|\sum_{n_1}^{n_2}\beta_n\right|<\zeta$$

for $n_2>n_1\geqslant\eta$; we also suppose that δP and η are not integers and that $\delta P<P-\eta$. Then

$$V_1^{(1)} = O\left(\sum_{m=1}^{\delta P}\frac{1}{m+1}\sum_{n=P-m+1}^{P+m}\frac{1}{n+1}\right) = O\left(\sum_{m=1}^{\delta P}\frac{1}{m}\cdot\frac{m}{P}\right) = O(\delta)$$

uniformly in P; $$V_1^{(2)} = O\left(\sum_{\delta P}^{P-\eta}\frac{\zeta}{m+1}\right) = O\left(\zeta\log\frac{1}{\delta}\right)$$

uniformly in P; and

$$V_1^{(3)} = O\left(\sum_{P-\eta}^{P}\frac{1}{m+1}\right) = O\left(\frac{\eta}{P}\right).$$

We can make $V_1^{(1)}$ and $V_1^{(2)}$ as small as we please by choice of δ, ζ, and η, and $V_1^{(3)}\to 0$ when δ, ζ, and η are fixed. Hence $V_1\to 0$.

The discussion of V_2 is similar. We write

$$V_2 = \Big(\sum_{m=P+1}^{P+\eta} + \sum_{P+\eta}^{\Delta P} + \sum_{\Delta P}^{\infty}\Big)\alpha_m \sum_{n=m-P}^{m+P} \beta_n = V_2^{(1)} + V_2^{(2)} + V_2^{(3)},$$

where $\Delta > 2$. Here

$$V_2^{(3)} = O\Big(\sum_{\Delta P}^{\infty} \frac{1}{m+1} \sum_{n=m-P}^{m+P} \frac{1}{n+1}\Big) = O\Big(P \sum_{\Delta P}^{\infty} \frac{1}{m^2}\Big) = O\Big(\frac{1}{\Delta}\Big),$$

uniformly in P; and $V_2^{(1)}$ and $V_2^{(2)}$ may be treated like $V_1^{(3)}$ and $V_1^{(2)}$ respectively. Hence

$$V_2 \to 0;$$

and this completes the proof.

It is plain after § 10.12 that there is a corresponding theorem for Laurent products, viz.: *if $a_m = O(|m|^{-1})$, $b_n = O(|n|^{-1})$, and A and B are unrestrictedly convergent, then the Laurent product C of A and B converges restrictedly to AB.* This assertion becomes false if either 'unrestrictedly' or 'restrictedly' stands in both hypotheses and conclusion. If

$$a_m = (m+2)^{ic-1}\{\log(m+2)\}^{-\rho}, \qquad b_n = (2-n)^{-ic-1}\{\log(2-n)\}^{-\rho},$$

where $c > 0$, $0 < \rho < \frac{1}{2}$, $m \geqslant 0$, $n \leqslant 0$, and a_m and b_n are 0 when $m < 0$, $n > 0$, then the hypotheses are satisfied with 'unrestrictedly', but $|C_{P,0}| > H(\log P)^{1-2\rho}$, so that C is not unrestrictedly convergent. If $a_0 = b_0 = 0$, $a_m = b_m = m^{-1}$ for $m \neq 0$, then the hypotheses are satisfied with 'restrictedly', and $A = B = 0$; but $c_0 = -\frac{1}{3}\pi^2$ and $c_p = c_{-p} = -2/p^2$ for $p \neq 0$, so that C converges (absolutely) to $-\pi^2 \neq AB$.

There is also a theorem corresponding to Theorem 173, viz.

THEOREM 195. *If A and B are restrictedly convergent,*

$$p > 1, \qquad q > 1, \qquad \sum |m|^{p-1}|a_m|^p < \infty, \qquad \sum |n|^{q-1}|b_n|^q < \infty,$$

then the Fourier product C converges to AB.

We leave the proof to the reader.

10.15. The analogue of Abel's theorem. It is natural to ask whether there is an analogue of Abel's Theorem 162, i.e.

(1) whether the unrestricted convergence of A, B, and their Laurent product C necessarily imply $C = AB$;†

(2) whether the restricted convergence of A and B, and the convergence of their Fourier product C implies $C = AB$.

Miss S. M. Edmonds, however, has constructed an example which shows that the answer to both questions is negative.‡ In this

$$a_m = m^{-\frac{3}{4}} \sin \pi\sqrt{m} \quad (m > 0), \qquad a_m = 0 \quad (m \leqslant 0), \qquad b_n = a_{-n},$$

† The second example of § 10.14 shows that, when A and B are only restrictedly convergent, their Laurent product C may converge, even absolutely, to a sum different from AB.

‡ Miss Edmonds considers only Laurent products, but the assertion about Fourier products is a simple corollary.

so that $A = B = \sum^{+} m^{-\frac{3}{4}} \sin \pi\sqrt{m}$; the Laurent product converges to A^2+2 and the Fourier product to A^2+1.

There is no very simple analogue of Cesàro's theorem, even when $r = s = 0$.

NOTES ON CHAPTER X

§ 10.2. Theorem 160 was first stated explicitly and proved satisfactorily by Cauchy, l.c. under § 1.1, 147.

Theorem 161 was proved by Mertens, *JM*, 79 (1875), 182–4. The negative remark which follows the proof is due to Schur (l.c. under § 3.2). I have stated the proof in a way suggested to me by Miss S. M. Edmonds, and arranged the proofs of Theorems 167, 190, and 191 on similar lines.

Theorem 162 is contained in Abel's memoir on the binomial theorem, *JM*, 1 (1826), 311–39 (317–18) [*Œuvres* (1), ed. 2 (1881), 219–50 (226)].

§ 10.3. Cesàro, *BSM* (2), 14 (1890), 114–20, proved Theorem 164 for integral r and s. The extension was made independently by Knopp and Chapman (l.c. under § 5.5.).

Theorem 167 was proved, for integral r, by Hardy and Littlewood, *PLMS* (2) 11 (1912), 411–78 (Theorem 35), and for general r by Hardy and Riesz, 65 (where it is extended to Dirichlet multiplication). The theorem for integral r is included in a more general theorem published a little before by Fekete, *MTE*, 29 (1911), 719–26, to the effect that if r and s are integers, A is absolutely summable (C, r) and B summable (C, s), then C is summable $(C, r+s)$. This in its turn was extended to general r and s by Kogbetliantz, *BSM* (2), 49 (1925), 234–56: see also Winn, *PEMS* (2), 3 (1933), 173–8. For the notion of absolute summability see the note on §§ 6.5–6.

§ 10.4. Theorem 169 was first proved by Hardy, *PLMS* (2), 6 (1908), 410–23, and has since been generalized, and the proof simplified, by a number of writers.

Theorem 170 is due to Neder, ibid., 23 (1923), 172–84 (except that Neder has $\eta = \zeta = \frac{1}{2}x$). The proof here follows Hardy, *PCPS*, 40 (1944), 251–2. Intermediate theorems, and generalizations in various directions, will be found in

Hardy, *PLMS* (2), 10 (1912), 396–405, and *JLMS*, 2 (1927), 169–71;

Rosenblatt, *BAP* (1913), 603–31, and *DMV*, 23 (1914), 80–4; Landau, *DMV*, 29 (1920), 238;

Broderick, *PLMS* (2), 19 (1921), 57–74, and 22 (1923), 468–82.

Some of the generalizations in these papers concern Dirichlet multiplication (§ 10.11).

In the first of his two papers Rosenblatt proves that if $r > 0$, $s > 0$, A is summable (C, r) and B summable (C, s), and

$$A_m^{r-1} = O(m^{r-1}), \qquad B_n^{s-1} = O(n^{s-1}),$$

then C is summable $(C, r+s)$. This result reduces to that of Theorem 169 when $r = s = 0$ and we interpret A_m^{-1} and B_n^{-1} as a_m and b_n. But this (as is suggested by the proof of Theorem 169 in the text) is not the best result. For A, being bounded $(C, r-1)$ and summable (C, r), is summable $(C, r-1+\delta)$, by Theorem 70; and similarly B is summable $(C, s-1+\delta)$. Hence, by Theorem 164, C is summable $(C, r+s-1+2\delta)$, i.e. by all means of order greater than $r+s-1$.

Hardy and Littlewood, l.c. under § 10.3 (464–6), show that if α and β are any numbers less than 1, there are convergent series A and B, with $a_m = O(m^{-\alpha})$, $b_n = O(n^{-\beta})$, whose product is divergent.

§ 10.5. Theorem 173 (with the generalization to Dirichlet multiplication) is proved by Hardy and Littlewood, *MM*, 43 (1914), 134–47 (137).

§ 10.6. Pringsheim, *MA*, 21 (1883), 327–78 (360–71). The proof here, which is simpler than Pringsheim's, was given by Hardy in the first paper quoted under § 10.4. See also Bromwich, 94–5.

§ 10.7. Theorems of this kind were first considered by Rajchmann, *Comptes rendus Soc. Sc. de Varsovie*, 11 (1918), 115–52: see Zygmund, *MZ*, 24 (1926), 47–104 (especially 48–65). We have sharpened the conditions. Rajchmann and Zygmund consider series infinite in both directions: see § 10.14.

§ 10.8. Bohr, *Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandlinger* (1908), 213–32, proved Theorems 178, 179, and 180, and the case $r = s = 0$ of Theorem 181. Chapman, l.c. under § 5.5, proved Theorem 181 generally.

§ 10.9. Knopp, *MZ*, 18 (1923), 125–56 (130–1), proves a theorem which, when combined with his theorems of § 8.3, gives the substance of the theorems here.

§ 10.10. Borel, 131–5, proves that the product of two series absolutely summable in his sense (i.e. regularly summable in the sense of § 8.6) is absolutely summable. Hardy, *QJM*, 35 (1903), 22–66, proves Theorem 189 and a part of Theorem 188, viz. that C is summable if A is absolutely summable in Borel's sense.

Theorem 187 is due in substance to Doetsch, *Dissertation*, Göttingen, 1920. Doetsch works in terms of the exponential definition, saying that A is summable (B, k) if $e^{-x}A(x) \to A\ (\mathrm{C}, k)$, and proves

$$\sum a_m = A\ (\mathrm{B}, r).\sum b_n = B\ (\mathrm{B}, s) \to \sum c_p = AB\ (\mathrm{B}, r+s+1).$$

In particular the summability (B) of A and B implies the summability (B, 1) of C. It is easy to prove that the assertions

$$a_0+a_1+... = A\ \ (\mathrm{B}, 1), \qquad a_1+a_2+... = A-a_0\ \ (\mathrm{B}';\ \mathrm{C}, 1)$$

are equivalent, and to deduce that Theorem 187 is equivalent to the case $r = s = 0$ of Doetsch's theorem.

Sannia, *RP*, 42 (1917), 303–22, generalizes the definitions differently, but his conclusions concerning multiplication are incorrect.

§ 10.11. Theorem 190 is due to Stieltjes, *NA* (3), 6 (1887), 210–15. Many examples of the use of the theorem will be found in Landau, *Handbuch*, 673 et seq., and in Ramanujan, *TCPS*, 22 (1918), 259–76 (*Collected papers*, no. 21).

For fuller information concerning Dirichlet multiplication see Hardy and Riesz, ch. 8; Landau, *RP*, 24 (1907), 81–160, and *Handbuch*, 750–67; and the papers quoted in the note on § 10.4.

A particularly striking theorem is that when $\lambda_m = \log m, \mu_n = \log n, \nu_p = \log p$, the convergence of A and B implies that of $\sum p^{-\frac{1}{2}}c_p$. This was stated without proof by Stieltjes, and proved by Landau, *l.c. supra*. Landau, *RP*, 26 (1908), 169–302 (265–6), proved that $\sum p^{-s}c_p$ is not necessarily convergent for all positive s, and Bohr, *WS*, 119 (1910), 1391–7, that the index $\frac{1}{2}$ cannot be replaced by any smaller number.

§§ 10.12–13. The problem was considered first by Chapman, *QJM*, 44 (1913), 219–33 (for Laurent multiplication). He proves Theorem 191.

§ 10.14. The first theorems of the type of Theorem 193 were those of Rajchmann and Zygmund: see the note on § 10.7. The other theorems of this section are referred to by S. M. Edmonds, *l.c. infra*, but have not been published before.

§ 10.15. S. M. Edmonds, *JLMS*, 17 (1942), 65–70.

XI

HAUSDORFF MEANS

11.1. The transformation δ. In this chapter we shall be concerned with a class of transformations which includes a number of those studied in earlier chapters, and in particular those of Cesàro, Hölder, and Euler. The theory depends upon the properties of the special transformation

$$t_m = \Delta^m s_0 = \sum_{n=0}^{m} (-1)^n \binom{m}{n} s_n. \tag{11.1.1}$$

We shall denote this transformation by

$$t = \delta s, \tag{11.1.2}$$

and the matrix

$$\begin{pmatrix} 1 & 0 & 0 & . & . & . \\ 1 & -1 & 0 & . & . & . \\ 1 & -2 & 1 & . & . & . \\ . & . & . & . & . & . \end{pmatrix} \tag{11.1.3}$$

associated with it by $|\delta|$.

THEOREM 196. *δ is its own reciprocal: if $t = \delta s$, then $s = \delta t$.*

Thus $\delta\delta = \mathrm{I}$, where I is the identity $t_m = s_m$.

For, if t_m is defined by (11.1.1), then

$$\Delta^m t_0 = \sum_{n=0}^{m} (-1)^n \binom{m}{n} t_n = \sum_{n=0}^{m} (-1)^n \binom{m}{n} \sum_{p=0}^{n} (-1)^p \binom{n}{p} s_p$$

$$= \sum_{p=0}^{m} (-1)^p s_p \sum_{n=p}^{m} (-1)^n \binom{m}{n}\binom{n}{p} = \sum_{p=0}^{m} (-1)^p \binom{m}{p} s_p \sum_{n=p}^{m} (-1)^n \binom{m-p}{n-p}$$

$$= \sum_{p=0}^{m} \binom{m}{p} s_p \sum_{q=0}^{m-p} (-1)^q \binom{m-p}{q} = s_m,$$

since

$$\binom{m}{n}\binom{n}{p} = \binom{m}{p}\binom{m-p}{n-p} \quad (0 \leqslant p \leqslant n \leqslant m)$$

and the inner sum in the last line is 1 if $p = m$ and 0 otherwise.†

It follows from the formulae of § 1.3 (4)‡ that if

$$x = -\frac{y}{1-y}, \qquad y = -\frac{x}{1-x}, \qquad s(x) = \sum s_n x^n, \qquad t(x) = \sum t_n x^n,$$

then

$$(1-x)s(x) = t(y).$$

† Symbolically, $t_n = (1-E)^n s_0$, $\Delta^m t_0 = \{1-(1-E)\}^m s_0 = E^m s_0 = s_m$.

‡ Changing the sign of x and replacing a_n and b_n by $(-1)^n s_n$ and t_n.

Theorem 196 shows that this implies $(1-x)t(x) = s(y)$, as may be verified directly. We have also seen in § 9.6 that if

$$S(x) = \sum \frac{s_n x^n}{n!}, \qquad T(x) = \sum \frac{t_n x^n}{n!},$$

then $e^{-x}S(x) = T(-x)$, and Theorem 196 shows that this is equivalent to $e^{-x}T(x) = S(-x)$ (as again may be verified directly).

11.2. Expression of the (E, q) and (C, 1) transformations in terms of δ. The (E, q) mean of s_n was defined by

$$t_m = \frac{1}{(q+1)^m} \sum_{n=0}^{m} \binom{m}{n} q^{m-n} s_n,$$

and (8.3.5) shows that

$$\Delta^n t_0 = (q+1)^{-n} \Delta^n s_0.$$

Hence, if we write $\Delta^n s_0 = u_n$, $\Delta^n t_0 = v_n$, and denote the diagonal transformation $t'_m = \mu_m s_m$ by μ, then we have $t = \delta v$, $v = \mu u$, $u = \delta s$, and so

(11.2.1) $$t = \lambda s,$$

where

(11.2.2) $$\lambda = \delta\mu\delta,$$

(11.2.3) $$\mu_n = (q+1)^{-n}.$$

Next, if t_m is the (C, 1) mean of s_n, so that

$$t_m = \frac{1}{m+1} \sum_{n=0}^{m} s_n,$$

then $$\Delta^n t_0 = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{1}{k+1} \sum_{l=0}^{k} s_l = \sum_{l=0}^{n} \phi_l s_l,$$

where $$\phi_l = \sum_{k=l}^{n} (-1)^k \binom{n}{k} \frac{1}{k+1}.$$

But this sum, written from $l = n$ downwards, is

$$\frac{(-1)^n}{n+1}\left\{1 - \binom{n}{1}\frac{n+1}{n} + \binom{n}{2}\frac{n+1}{n-1} - \dots\right\} = \frac{(-1)^n}{n+1}\left\{1 - \binom{n+1}{1} + \binom{n+1}{2} - \dots\right\},$$

the series being continued for $n-l+1$ terms; and hence

$$\phi_l = \frac{(-1)^l}{n+1}\binom{n}{l}.\dagger$$

† The sum of the first p coefficients in the expansion of $(1-x)^{n+1}$ is the pth coefficient in that of $(1-x)^n$.

Thus $$\Delta^n t_0 = \frac{1}{n+1}\sum_{l=0}^{n}(-1)^l \binom{n}{l} s_l = \frac{1}{n+1}\Delta^n s_0;$$

and it follows again that the transformation can be expressed in the form (11.2.2), with

(11.2.4) $$\mu_n = (n+1)^{-1}.$$

11.3. Hausdorff's general transformation. We call the transformation

(11.3.1) $$t = (\delta\mu\delta)s = \lambda s,$$

where μ is any diagonal transformation, a Hausdorff or $\mathfrak{H}$ transformation, and its matrix an $\mathfrak{H}$ matrix. Thus the (E, q) and $(\mathrm{C}, 1)$ transformations are $\mathfrak{H}$ transformations. We shall use $\mathfrak{H}$, or $(\mathfrak{H}, \mu)$ for the transformation, $|\mathfrak{H}|$ or $|\mathfrak{H}, \mu|$ for its matrix.

If $\mathfrak{H} = \delta\mu\delta$, $\mathfrak{H}' = \delta\mu'\delta$, then

$$\mathfrak{H}\mathfrak{H}' = \delta\mu\delta\delta\mu'\delta = \delta\mu\mu'\delta = \delta\mu'\mu\delta = \mathfrak{H}'\mathfrak{H}.$$

Thus

THEOREM 197. *Any two $\mathfrak{H}$ transformations are commutable.*

Conversely, suppose that $\gamma = \delta\mu\delta$ is a given $\mathfrak{H}$ transformation; that the numbers μ_n are all different; and that λ is any transformation commutable with γ. If $\varpi = \delta\lambda\delta$, then $\lambda = \delta\varpi\delta$. Also $\mu = \delta\gamma\delta$. Hence

$$\varpi\mu = \delta\lambda\delta\delta\gamma\delta = \delta\lambda\gamma\delta, \qquad \mu\varpi = \delta\gamma\delta\delta\lambda\delta = \delta\gamma\lambda\delta;$$

and $\lambda\gamma = \gamma\lambda$, by hypothesis, so that

(11.3.2) $$\varpi\mu = \mu\varpi.$$

If ϖ is $$t_m = \sum c_{m,n} s_n,$$

then (11.3.2) implies

$$\sum c_{m,n}\mu_n s_n = \mu_m \sum c_{m,n} s_n$$

for all s_n; and since $\mu_m \neq \mu_n$ when $m \neq n$, this implies that $c_{m,n} = 0$ when $m \neq n$. Hence ϖ is a diagonal transformation, and $\lambda = \delta\varpi\delta$ is an $\mathfrak{H}$ transformation.

The condition on γ is satisfied by the $(\mathrm{C}, 1)$ transformation. Hence

THEOREM 198. *The class of $\mathfrak{H}$ transformations is that of transformations commutable with the* $(\mathrm{C}, 1)$ *transformation (or any other $\mathfrak{H}$ transformation all of whose μ_n differ).*

It is easy to determine the coefficients in any $\mathfrak{H}$ transformation in terms of its μ_n. For, using u_n and v_n as in § 11.2, we have $t = \delta v$, i.e.

$$t_m = \sum_{n=0}^{m} (-1)^n \binom{m}{n} \Delta^n t_0 = \sum_{n=0}^{m} (-1)^n \binom{m}{n} \mu_n \Delta^n s_0$$

$$= \sum_{n=0}^{m} (-1)^n \binom{m}{n} \mu_n \sum_{p=0}^{n} (-1)^p \binom{n}{p} s_p = \sum_{p=0}^{m} \phi_{m,p} s_p,$$

where

$$\phi_{m,p} = (-1)^p \sum_{n=p}^{m} (-1)^n \binom{m}{n}\binom{n}{p} \mu_n = (-1)^p \binom{m}{p} \sum_{n=p}^{m} (-1)^n \binom{m-p}{n-p} \mu_n$$

$$= \binom{m}{p} \sum_{k=0}^{m-p} (-1)^k \binom{m-p}{k} \mu_{p+k} = \binom{m}{p} \Delta^{m-p} \mu_p;$$

and so, writing n again for p,

$$t_m = \sum_{n=0}^{m} \binom{m}{n} \Delta^{m-n} \mu_n \,.\, s_n.$$

THEOREM 199. *The general $\mathfrak{H}$ transformation is*

(11.3.3) $$t_m = \sum \lambda_{m,n} s_n,$$

where

(11.3.4) $$\lambda_{m,n} = \binom{m}{n} \Delta^{m-n} \mu_n \quad (n \leqslant m), \qquad 0 \quad (n > m).$$

We shall write

(11.3.5) $$\mu_{n,p} = \Delta^p \mu_n,$$

so that

(11.3.6) $$\lambda_{m,n} = \binom{m}{n} \mu_{n,m-n} \quad (0 \leqslant n \leqslant m).$$

11.4. The general Hölder and Cesàro transformations as $\mathfrak{H}$ transformations. We denote the (H, k) and (C, k) transformations by $H^{(k)}$ and $C^{(k)}$, as in § 5.9, and write H and C for $H^{(1)}$ and $C^{(1)}$: $H^{(k)}$ is so far defined for $k = 0, 1, 2, \ldots$ only, and $H^{(k)} = H^k$, i.e. the result of k repetitions of H.

If $\lambda = \delta\mu\delta$, $\lambda' = \delta\mu'\delta$ then

$$\lambda\lambda' = \delta\mu\delta\delta\mu'\delta = \delta\mu\mu'\delta.$$

It follows that $H^{(k)}$ is the $\mathfrak{H}$ transformation corresponding to

$$\mu_n = (n+1)^{-k}.$$

On the other hand, it is not obvious that $C^{(k)}$ is an $\mathfrak{H}$ transformation, even when k is an integer. We proceed to show that it is one, and to determine the corresponding μ_n.

In this case

$$t_m = \binom{m+k}{k}^{-1} \sum_{n=0}^{m} \binom{m-n+k-1}{k-1} s_n,$$

$$\Delta^n t_0 = \sum_{p=0}^{n} (-1)^p \binom{n}{p} t_p$$

$$= \sum_{p=0}^{n} (-1)^p \binom{p+k}{k}^{-1} \binom{n}{p} \sum_{q=0}^{p} \binom{p-q+k-1}{k-1} s_q = \sum_{q=0}^{n} \phi_{n,q} s_q,$$

where

$$\phi_{n,q} = \sum_{p=q}^{n} (-1)^p \binom{p+k}{k}^{-1} \binom{n}{p} \binom{p-q+k-1}{k-1}$$

$$= k\Gamma(n+1) \sum_{p=q}^{n} (-1)^p \frac{\Gamma(p-q+k)}{\Gamma(n-p+1)\Gamma(p-q+1)\Gamma(p+k+1)}$$

$$= (-1)^n k\Gamma(n+1) \frac{\Gamma(n-q+k)}{\Gamma(n-q+1)\Gamma(n+k+1)} \left\{1 - \frac{(n-q)(n+k)}{1(n-q+k-1)} + \right.$$

$$\left. + \frac{(n-q)(n-q-1)(n+k)(n+k-1)}{1.2(n-q+k-1)(n-q+k-2)} - \ldots\right\}$$

$$= (-1)^n k\Gamma(n+1) \frac{\Gamma(n-q+k)}{\Gamma(n-q+1)\Gamma(n+k+1)} \frac{\Gamma(-n+q-k+1)\Gamma(n+1)}{\Gamma(-k+1)\Gamma(q+1)}$$

$$= (-1)^q \binom{n+k}{k}^{-1} \binom{n}{q}.\dagger$$

Hence $$\Delta^n t_0 = \binom{n+k}{k}^{-1} \sum_{q=0}^{n} (-1)^q \binom{n}{q} s_q = \binom{n+k}{k}^{-1} \Delta^n s_0,\ddagger$$

and the transformation is an $\mathfrak{H}$ transformation with

$$\mu_n = \binom{n+k}{k}^{-1}.$$

Thus

Theorem 200. *The* $H^{(k)}$ *and* $C^{(k)}$ *transformations are* $\mathfrak{H}$ *transformations, with*

(11.4.1) $$\mu_n = \frac{1}{(n+1)^k}, \qquad \mu_n = \binom{n+k}{k}^{-1}.$$

† By Gauss's formula for the sum of the hypergeometric series $F(\alpha, \beta; \gamma; 1)$, and the equation $\Gamma(x)\Gamma(1-x) = \pi \operatorname{cosec} x\pi$. It is convenient to suppose k non-integral in making the calculations.

‡ This is (9.6.9): the proof there was less direct.

Theorem 200 has been proved for all k for which we have defined the Hölder and Cesàro means, i.e. for $k = 0, 1, 2,...$ in the first case and $k > -1$ in the second. It leads us naturally to define $H^{(k)}$, for non-integral k, as the $\mathfrak{H}$ transformation with $\mu_n = (n+1)^{-k}$. We shall see later (§ 11.11) that the two systems of means are then equivalent for all $k > -1$.

11.5. Conditions for the regularity of a real Hausdorff transformation. In order that the transformation

(11.5.1) $$t_m = \sum c_{m,n} s_n$$

should be regular, it is necessary and sufficient, after Theorem 2, (1) that

(11.5.2) $$\gamma_m = \sum_n |c_{m,n}| < K,$$

where K is independent of m; (2) that

(11.5.3) $$c_{m,n} \to 0$$

for every n, when $m \to \infty$; and (3) that

(11.5.4) $$c_m = \sum_n c_{m,n} \to 1$$

when $m \to \infty$. We have now to interpret these conditions, for an $\mathfrak{H}$ transformation, in terms of μ_n. We suppose μ_n real.

If $s_n = 1$ for all n, then $u_n = \Delta^n s_0$ and $v_n = \mu_n u_n$ are 1 and μ_0 respectively for $n = 0$ and 0 for $n > 0$, so that $t_n = \Delta^n v_0 = \mu_0$ for all n. Hence

(11.5.5) $$\sum_{n=0}^{m} \binom{m}{n} \mu_{n,m-n} = \mu_0$$

(as may, of course, be verified directly),† and (11.5.4) reduces to

(11.5.6) $$\mu_0 = 1.$$

Thus the conditions (11.5.2), (11.5.3), and (11.5.4) are

(11.5.7) $$M_m = \sum_{n=0}^{m} \binom{m}{n} |\mu_{n,m-n}| < K,$$

(11.5.8) $$\binom{m}{n} \mu_{n,m-n} \to 0 \quad (n = 0, 1,...),$$

(11.5.9) $$\mu_0 = 1.$$

We proceed to analyse the meaning of (11.5.7), and to show that, when it is satisfied, (11.5.8) may be replaced by a simpler condition.

† For example,

$$\sum \binom{m}{n} \Delta^{m-n} \mu_n = \sum \binom{m}{n} \Delta^{m-n} E^n \mu_0 = (\Delta + E)^m \mu_0 = \mu_0.$$

11.6. Totally monotone sequences. In the sections which follow we shall be concerned with real sequences only. A sequence (ϕ_n) will be said to be *totally monotone*† if

(11.6.1) $$\Delta^p \phi_n \geqslant 0$$

for $n = 0, 1,..., p = 0, 1,....$ Thus

$$\Delta^p \frac{1}{n+1} = \frac{p!}{(n+1)(n+2)...(n+p+1)} > 0,$$

so that the μ_n of the (C, 1) transformation is totally monotone. If μ_n is totally monotone, then $\mu_{n,m-n} = \Delta^{m-n}\mu_n \geqslant 0$, and

(11.6.2) $$M_m = \sum \binom{m}{n} \mu_{n,m-n} = \mu_0,$$

by (11.5.5), so that (11.5.7) is certainly satisfied. Analogy with the theory of functions or sequences of bounded variation (an analogy which we shall find to be closer than appears at first sight) then suggests the truth of the following theorem.

THEOREM 201. *In order that a real μ_n should satisfy* (11.5.7), *it is necessary and sufficient that*

(11.6.3) $$\mu_n = \alpha_n - \beta_n,$$

where α_n and β_n are totally monotone.

It is obvious that the condition is sufficient, by (11.5.5), and we have to prove it necessary.

We write ($Eu_n = u_{n+1}$ and)

$$E_1 u_{n,p} = u_{n+1,p}, \qquad E_2 u_{n,p} = u_{n,p+1}.$$

Then

(11.6.4) $$\mu_{n,p} = \Delta^p \mu_n = (E+\Delta)\Delta^p \mu_n = \Delta^p \mu_{n+1} + \Delta^{p+1}\mu_n$$
$$= \mu_{n+1,p} + \mu_{n,p+1} = (E_1+E_2)\mu_{n,p}$$

and

(11.6.5) $$|\mu_{n,p}| \leqslant |\mu_{n+1,p}| + |\mu_{n,p+1}| = (E_1+E_2)|\mu_{n,p}|.$$

If

(11.6.6) $$\mu_{n,p,m} = \sum_{r=0}^{m} \binom{m}{r} \mu_{n+r,p+m-r}, \qquad \mu^*_{n,p,m} = \sum_{r=0}^{m} \binom{m}{r} |\mu_{n+r,p+m-r}|,$$

then

(11.6.7) $$\mu_{n,p,m} = \sum_{r=0}^{m} \binom{m}{r} E_1^r E_2^{m-r} \mu_{n,p} = (E_1+E_2)^m \mu_{n,p} = \mu_{n,p},$$

(11.6.8) $$\mu^*_{n,p,m} = \sum_{r=0}^{m} \binom{m}{r} E_1^r E_2^{m-r} |\mu_{n,p}| = (E_1+E_2)^m |\mu_{n,p}| \geqslant |\mu_{n,p}|,$$

† More properly, perhaps, 'totally decreasing'. A sequence such as $(\phi_n) = (e^n)$, in which $\Delta^p\phi_n$ has the sign $(-1)^p$, might be called 'totally increasing'.

by (11.6.4) and (11.6.5). Also

(11.6.9) $$\mu^*_{0,0,m} = M_m < K,$$

by (11.5.7). Now

$$\mu^*_{n,p,m} = (E_1+E_2)^m|\mu_{n,p}| \leqslant (E_1+E_2)^{m+1}|\mu_{n,p}| = \mu^*_{n,p,m+1},$$

by (11.6.5), so that $\mu^*_{n,p,m}$ increases with m. Also

$$\begin{aligned}\mu^*_{n,p,m} &= (E_1+E_2)^m|\mu_{n,p}| = (E_1+E_2)^m E_1^n E_2^p|\mu_{0,0}|\\ &\leqslant \binom{n+p}{p}(E_1+E_2)^m E_1^n E_2^p|\mu_{0,0}|\\ &\leqslant (E_1+E_2)^m \sum_{r=0}^{n+p}\binom{n+p}{r} E_1^{n+p-r}E_2^r|\mu_{0,0}|\\ &= (E_1+E_2)^m(E_1+E_2)^{n+p}|\mu_{0,0}| = (E_1+E_2)^{n+p+m}|\mu_{0,0}|\\ &= \mu^*_{0,0,n+p+m} = M_{n+p+m} < K,\end{aligned}$$

by (11.5.7). Hence

(11.6.10) $$\mu^*_{n,p,m} \to \lim_{m\to\infty} \mu^*_{n,p,m} = \mu^*_{n,p},$$

say, when $m \to \infty$. Also $|\mu_{n,p}| = |\mu_{n,p,m}| \leqslant \mu^*_{n,p,m}$, by (11.6.7) and (11.6.8), and so

(11.6.11) $$\mu_{n,p} \leqslant \mu^*_{n,p}.$$

In particular

(11.6.12) $$|\mu_n| = |\mu_{n,0}| \leqslant \mu^*_{n,0} = \mu^*_n,$$

say.

Next $$\mu^*_{n,p,m+1} = (E_1+E_2)^{m+1}|\mu_{n,p}| = (E_1+E_2)\mu^*_{n,p,m},$$

by (11.6.8), and so

$$\mu^*_{n,p+1,m} = \mu^*_{n,p,m+1} - \mu^*_{n+1,p,m}.$$

Hence, making $m \to \infty$,

$$\mu^*_{n,p+1} = \mu^*_{n,p} - \mu^*_{n+1,p} = \Delta\mu^*_{n,p},$$

and so $\mu^*_{n,p} = \Delta^p\mu^*_{n,0} = \Delta^p\mu^*_n$. Thus

$$|\Delta^p\mu_n| = |\mu_{n,p}| \leqslant \mu^*_{n,p} = \Delta^p\mu^*_n.$$

Hence, finally, if we write

$$\alpha_n = \tfrac{1}{2}(\mu^*_n+\mu_n), \qquad \beta_n = \tfrac{1}{2}(\mu^*_n-\mu_n),$$

then $$\mu_n = \alpha_n - \beta_n, \qquad \Delta^p\alpha_n \geqslant 0, \qquad \Delta^p\beta_n \geqslant 0,$$

which proves the theorem.

11.7. Final form of the conditions for regularity. We now assume that condition (11.5.7) is satisfied, and use it to simplify condition (11.5.8). We shall prove that (11.5.7) implies (11.5.8) for $n > 0$. When $n = 0$, (11.5.8) is

(11.7.1) $$\Delta^m \mu_0 \to 0;$$

and this, which is not a consequence of (11.5.7), must be kept as a separate condition.

A sequence which satisfies (11.5.7) is the difference of two totally monotone sequences. It is therefore sufficient to prove that

(11.7.2) $$\lambda_{m,n} = \binom{m}{n} \mu_{n,m-n} \to 0$$

when $n > 0$ and (μ_n) is totally monotone, so that $\lambda_{m,n} \geqslant 0$. It follows from (11.3.5) and (11.3.6) that

$$\mu_{n,m-n} = \mu_{n,m-n+1} + \mu_{n+1,m-n},$$

$$(m+1)\lambda_{m,n} = (m-n+1)\lambda_{m+1,n} + (n+1)\lambda_{m+1,n+1},$$

or $$(m+1)(\lambda_{m,n} - \lambda_{m+1,n}) = (n+1)\lambda_{m+1,n+1} - n\lambda_{m+1,n}.$$

Summing with respect to n, and writing

$$\Lambda_{m,n} = \lambda_{m,0} + \lambda_{m,1} + \ldots + \lambda_{m,n},$$

we obtain

(11.7.3) $$(m+1)(\Lambda_{m,n} - \Lambda_{m+1,n}) = (n+1)\lambda_{m+1,n+1} \geqslant 0.$$

Hence $\Lambda_{m,n}$ decreases as m increases, and tends to a limit when $m \to \infty$; and therefore $\lambda_{m,n} = \Lambda_{m,n} - \Lambda_{m,n-1}$ tends to a limit l_n. In particular $\lambda_{m,0} = \Lambda_{m,0} \to l_0$. Also, for $n \geqslant 0$,

$$\rho_m = \Lambda_{m,n} - \Lambda_{m+1,n} \sim \frac{n+1}{m+1} l_{n+1},$$

when $m \to \infty$, by (11.7.3), and $\sum \rho_m$ is convergent, so that $l_{n+1} = 0$. Hence

(11.7.4) $$\lambda_{m,0} \to l_0,$$ (11.7.5) $$\lambda_{m,n} \to 0 \quad (n > 0).$$

Thus (11.7.5), which is (11.5.8) for $n > 0$, is a consequence of (11.5.7). But there is nothing to show that $l_0 = 0$, and this condition, which is (11.7.1), must be retained. The sequence $(1, 0, 0, \ldots)$ is totally monotone, but here $\Delta^m \mu_0 = 1$ for all m.

We have thus proved

THEOREM 202. *In order that the transformation $(\mathfrak{H}, \mu)$ should be regular, it is necessary and sufficient that (μ_n) should be the difference of two totally monotone sequences, that*

$$\Delta^m \mu_0 \to 0, \tag{11.7.6}$$

and that

$$\mu_0 = 1. \tag{11.7.7}$$

It is important to notice what follows from the main condition (11.5.7) alone, without the two subsidiary 'normalizing' conditions. We have then

$$c_{m,0} = \lambda_{m,0} \to l_0, \qquad c_{m,n} = \lambda_{m,n} \to 0 \quad (n > 0),$$

and $\sum c_{m,n} = \mu_0$ for all m. Thus the conditions of Theorem 1 are satisfied, with

$$\delta_0 = l_0, \qquad \delta_n = 0 \quad (n > 0), \qquad \delta = \mu_0.$$

The transformation preserves convergence (belongs to $\mathfrak{T}_c$), and

$$t_m \to \mu_0 s + l_0(s_0 - s)$$

whenever $s_n \to s$.

The condition $l_0 = 0$ excludes, for example, the sequence $(1, 0, 0, \ldots)$, while $\mu_0 = 1$ excludes $(2, 2, 2, \ldots)$. Both exclude $(2, 1, 1, \ldots)$. The transformation defined by the second of these sequences becomes regular (in fact the identity) when μ_n is divided by 2. In the third, $\Delta^m \mu_0$ is 2 for $m = 0$ and 1 for $m > 0$; the transformation becomes regular if μ_0 is decreased by 1. The significance of these supplementary conditions will become clearer when we have proved Hausdorff's theorem about the integral representation of μ_n.

11.8. Moment constants. We call

$$\mu_n = \int_0^1 x^n \, d\chi,\dagger \tag{11.8.1}$$

where $\chi = \chi(x)$ is a real function of bounded variation in $0 \leqslant x \leqslant 1$, the *moment constant*, of rank n, of χ. We may suppose without loss of generality that

$$\chi(0) = 0. \tag{11.8.2}$$

If also

$$\chi(1) = 1 \tag{11.8.3}$$

and

$$\chi(+0) = \chi(0) = 0 \tag{11.8.4}$$

so that $\chi(x)$ is continuous at the origin, then we shall call μ_n a *regular* moment constant.

† The function x^0 is defined at $x = 0$ so as to be continuous. Thus $\mu_0 = \int_0^1 d\chi$.

If $\chi(x)$ increases with x, then

$$\mu_{n,p} = \Delta^p \mu_n = \int x^n (1-x)^p \, d\chi \geqslant 0,\dagger$$

so that μ_n is totally monotone. Generally, if $P(x)$ and $N(x)$ are the positive and negative variations of $\chi(t)$ in $(0,x)$, then $\chi(x) = P(x)-N(x)$ and

$$\mu_n = \int x^n \, dP - \int x^n \, dN = \alpha_n - \beta_n,$$

where (α_n) and (β_n) are totally monotone.

The function $\chi(x)$ may have an enumerable set of discontinuities, and the value of the integral (11.8.1) is not affected by any change in the value of $\chi(x)$ at a point of discontinuity inside $(0,1)$. In particular we may suppose that

(11.8.5) $$\chi(x) = \tfrac{1}{2}\{\chi(x-0)+\chi(x+0)\}$$

for $0 < x < 1$, in which case we shall say that all discontinuities of $\chi(x)$ are *normal*. The expression of μ_n as a moment constant is then if possible, unique.

This follows from

THEOREM 203. *If*

$$\mu_n = \int x^n \, d\chi_1 = \int x^n \, d\chi_2,$$

where χ_1 and χ_2 are functions of bounded variation, vanishing at the origin and with normal discontinuities, then $\chi_1 = \chi_2$ for all x.

It is sufficient to show that, if

(11.8.6) $$\int x^n \, d\chi = 0 \quad (n = 0, 1, 2, \ldots),$$

$\chi(0) = 0$, and $\chi(x)$ satisfies (11.8.5), then $\chi(x) = 0$ for all x. It follows from (11.8.6), with $n = 0$, that $\chi(1) = 0$. Hence, integrating by parts,

$$n \int x^{n-1} \chi(x) \, dx = 0 \quad (n = 1, 2, \ldots),$$

and so $\int x^n \chi(x) \, dx = 0$ for $n \geqslant 0$. And if we write

$$\psi(x) = \int_0^x \chi(t) \, dt,$$

and integrate again by parts, we obtain

(11.8.7) $$\int x^n \psi(x) \, dx = 0 \quad (n = 0, 1, 2, \ldots).$$

† Here, when the limits of an integral are not shown, they are 0 and 1.

Since $\psi(x)$ is continuous, there is a polynomial $Q(x)$ such that $|\psi-Q| < \epsilon$ for $0 \leqslant x \leqslant 1$. Then, by (11.8.7),

$$\int \psi^2\,dx = \int \psi Q\,dx + \int \psi(\psi-Q)\,dx = \int \psi(\psi-Q)\,dx \leqslant \epsilon \int |\psi|\,dx,$$

and so, since ϵ is arbitrary, $\int \psi^2\,dx = 0$. Hence $\psi = 0$ for all x; and hence $\chi = 0$ at all its points of continuity. Since these points are dense in $(0,1)$, and $\chi(x-0)$ and $\chi(x+0)$ exist for every x, it follows that $\chi(x-0) = \chi(x+0) = 0$; and therefore by (11.8.5), that $\chi(x) = 0$.

We now interpret the conditions (11.8.3) and (11.8.4) in terms of μ_n. First, it is plain that (11.8.3) is equivalent to $\mu_0 = 1$, i.e. to (11.7.7). Next

$$\Delta^m \alpha_0 = \int (1-x)^m\,dP, \qquad \Delta^m \beta_0 = \int (1-x)^m\,dN$$

are non-negative and decrease as m increases, so that

$$\Delta^m \alpha_0 \to a \geqslant 0, \qquad \Delta^m \beta_0 \to b \geqslant 0, \qquad \Delta^m \mu_0 \to a-b.$$

We can choose η so that $0 < \eta < 1$ and $P(\eta) < P(+0)+\epsilon$; and then

$$\Delta^m \alpha_0 \leqslant \int_0^\eta dP + (1-\eta)^m \int_\eta^1 dP \leqslant P(\eta)+(1-\eta)^m P(1) < P(+0)+2\epsilon$$

for sufficiently large m, so that $a \leqslant P(+0)$.

On the other hand,

$$\Delta^m \alpha_0 \geqslant (1-\eta)^m \int_0^\eta dP = (1-\eta)^m P(\eta) > P(+0)-\epsilon$$

for $\eta < \eta(\epsilon, m)$. Hence $\Delta^m \alpha_0 \geqslant P(+0)$, and so $a \geqslant P(+0)$.

Thus $a = P(+0)$, and similarly $b = N(+0)$. It follows that

$$\Delta^m \mu_0 \to P(+0)-N(+0) = \chi(+0);$$

in particular (11.8.4) is equivalent to (11.7.6).

Summing up, we have proved

THEOREM 204. *Any moment constant μ_n is the difference of two totally monotone sequences. The moment constant of an increasing χ is totally monotone.*

THEOREM 205. *In order that a moment constant μ_n should satisfy the conditions of Theorem 202, and so define a regular $\mathfrak{H}$ transformation $(\mathfrak{H}, \mu)$, it is necessary and sufficient that μ_n should be regular.*

11.9. Hausdorff's theorem. We now prove Hausdorff's fundamental theorem, which shows that the results of Theorem 204 are reversible.

THEOREM 206. *If (μ_n) is the difference of two totally monotone sequences (α_n) and (β_n), then μ_n is a moment constant.*

It is plainly sufficient, after § 11.8, to prove

THEOREM 207. *If (μ_n) is totally monotone, then $\mu_n = \int x^n\,d\chi$, where $\chi(x)$ is an increasing and bounded function of x.*

The proof depends upon an important general theorem of Helly: if $\chi_q(x)$ is a sequence of increasing functions of x, uniformly bounded for $0 \leqslant x \leqslant 1$, then there is a bounded increasing function $\chi(x)$, and a subsequence (q_i) of values of q, such that $\chi_q(x) \to \chi(x)$ when $q \to \infty$ through (q_i).

We define $\chi_q(x)$ by

$$\chi_q(0) = 0, \qquad \chi_q(x) = \sum_{0 \leqslant s \leqslant qx} \lambda_{q,s} = \sum_{0 \leqslant s \leqslant qx} \binom{q}{s} \mu_{s,q-s} \quad (0 < x \leqslant 1).$$

Then $\chi_q(x)$ increases with x; and (11.5.5) shows that $\chi_q(1) = \mu_0$, so that $\chi_q(x)$ is uniformly bounded.

It follows that

$$\mu_0 = \chi_q(1) - \chi_q(0) = \lim\{\chi_{q_i}(1) - \chi_{q_i}(0)\} = \chi(1) - \chi(0) = \int d\chi.$$

If $n > 0$, then

$$\begin{aligned} \mu_n = \mu_{n,0} &= \mu_{n,1} + \mu_{n+1,0} = \mu_{n,2} + 2\mu_{n+1,1} + \mu_{n+2,0} \\ &= \ldots = \sum_{k=0}^{q-n} \binom{q-n}{k} \mu_{n+k,q-n-k} \end{aligned}$$

for all $q \geqslant n$. This is

$$\sum_{k=0}^{q-n} \frac{(q-n)!}{k!(q-n-k)!} \frac{(n+k)!\,(q-n-k)!}{q!} \binom{q}{n+k} \mu_{n+k,q-n-k}$$

$$= \sum_{k=0}^{q-n} \frac{(q-n)!(n+k)!}{k!\,q!} \binom{q}{n+k} \mu_{n+k,q-n-k} = \sum_{s=n}^{q} \frac{(q-n)!\,s!}{q!(s-n)!} \binom{q}{s} \mu_{s,q-s}$$

$$= \sum_{s=n}^{q} \frac{s(s-1)\ldots(s-n+1)}{q(q-1)\ldots(q-n+1)} \binom{q}{s} \mu_{s,q-s} = \sum_{s=0}^{q} \frac{s(s-1)\ldots(s-n+1)}{q(q-1)\ldots(q-n+1)} \binom{q}{s} \mu_{s,q-s}$$

(the terms added being all zero).

We divide $(0,1)$ by points $x_0 = 0$, $x_1,\ldots$, $x_r = 1$, suppose q large enough to make $qx_1 > n$, and write

$$S_l^{(q)} = \sum_{qx_l < s \leqslant qx_{l+1}} \frac{s(s-1)\ldots(s-n+1)}{q(q-1)\ldots(q-n+1)} \binom{q}{s} \mu_{s,q-s},$$

so that

$$\mu_n = \sum_{l=0}^{r-1} S_l^{(q)}.$$

Then, since
$$\chi_q(x_1)-\chi_q(x_0) = \sum_{0\leqslant s\leqslant qx_1}\binom{q}{s}\mu_{s,q-s},$$
$$\chi_q(x_{l+1})-\chi_q(x_l) = \sum_{qx_l<s\leqslant qx_{l+1}}\binom{q}{s}\mu_{s,q-s}\quad (l>0),$$
we have

$$\text{(11.9.1)}\qquad \frac{x_l q(x_l q-1)\dots(x_l q-n+1)}{q(q-1)\dots(q-n+1)}\{\chi_q(x_{l+1})-\chi_q(x_l)\}$$
$$\leqslant S_l^{(q)} \leqslant \frac{x_{l+1} q(x_{l+1} q-1)\dots(x_{l+1} q-n+1)}{q(q-1)\dots(q-n+1)}\{\chi_q(x_{l+1})-\chi_q(x_l)\}.\dagger$$

Hence, first,
$$\mu_n = \sum_{l=0}^{r-1} S_l^{(q)} \leqslant \sum_{l=0}^{r-1}\frac{x_{l+1} q(x_{l+1} q-1)\dots(x_{l+1} q-n+1)}{q(q-1)\dots(q-n+1)}\{\chi_q(x_{l+1})-\chi_q(x_l)\}.$$

From this, making $q\to\infty$ through an appropriate sequence (q_i), we obtain
$$\mu_n \leqslant \varlimsup_{q=q_i\to\infty}\sum_{l=0}^{r-1} S_l^{(q)} \leqslant \sum_{l=0}^{r-1} x_{l+1}^n\{\chi(x_{l+1})-\chi(x_l)\}.$$

We can obtain a lower bound for μ_n similarly from the first of the inequalities (11.9.1); and so

$$\text{(11.9.2)}\qquad \sum_{l=0}^{r-1} x_l^n\{\chi(x_{l+1})-\chi(x_l)\} \leqslant \mu_n \leqslant \sum_{l=0}^{r-1} x_{l+1}^n\{\chi(x_{l+1})-\chi(x_l)\}.$$

But the Stieltjes integral $\int x^n\,d\chi$ is the common limit of the two sums in (11.9.2) when r tends to infinity and the largest interval (x_l, x_{l+1}) tends to 0; and therefore $\mu_n = \int x^n\,d\chi$. The integral is not affected by any change in the value of χ at its discontinuities inside $(0,1)$. We may suppose that they are selected so as to normalize the discontinuities.

We can now resume our results in

THEOREM 208. (i) *In order that* $(\mathfrak{H},\mu)$ *should be a regular* $\mathfrak{H}$ *transformation, it is necessary and sufficient that* μ_n *should be a regular moment constant.* (ii) *In order that* $(\mathfrak{H},\mu)$ *should be a convergence preserving transformation it is necessary and sufficient that* μ_n *should be a moment constant.*

In the general case the variation of $\chi_q(t)$ in $(0,x)$ is given by
$$V_q(0)=0,\qquad V_q(x) = \sum_{0\leqslant s\leqslant qx}\binom{q}{s}|\mu_{s,q-s}|\quad (0<x\leqslant 1);$$

† This is obvious if $l>0$. If $l=0$, then the first member in (11.9.1) is 0, while the terms of the second are 0 for $s\leqslant n-1$, and the remainder non-negative and not greater than the corresponding terms of the third, whose terms are all non-negative.

and the variation of $\chi(t) = \lim \chi_{q_i}(t)$ in $(0, x)$ is $V(x) = \lim V_{q_i}(x)$. For it is easily shown that, if $0 \leqslant a < b \leqslant 1$, then

$$\int_a^b |d\chi| \leqslant \underline{\lim} \int_a^b |d\chi_{q_i}|.$$

On the other hand,

$$\int_0^1 |d\chi_q| = \sum_{s=0}^{q} \binom{q}{s} |\mu_{s,q-s}| = \sum_{s=0}^{q} \binom{q}{s} \left| \int_0^1 x^s(1-x)^{q-s}\, d\chi \right|$$
$$\leqslant \int_0^1 \left\{ \sum_{s=0}^{q} \binom{q}{s} x^s(1-x)^{q-s} \right\} |d\chi| = \int_0^1 |d\chi|.$$

The functions V_q and V are derived from $|\mu_{s,q-s}|$ as χ_q and χ are derived from $\mu_{s,q-s}$. Thus V corresponds to the μ_n^* of § 11.6 as χ corresponds to μ_n, and

$$\mu_n^* = \int x^n\, dV = \int x^n\, dP + \int x^n\, dN = \alpha_n + \beta_n.$$

Any expression $\chi = \theta - \phi$ of χ as the difference of two increasing functions corresponds to an expression $\mu_n = \rho_n - \sigma_n$ of μ_n as the difference of two totally monotone sequences. The decomposition $\chi = P - N$ is the 'least' in the sense that $\theta = P + \omega$, $\phi = N + \omega$, where ω is an increasing function; and the decomposition $\mu_n = \alpha_n - \beta_n$ is the 'least' in the sense that the components ρ_n and σ_n of any other decomposition are of the forms $\rho_n = \alpha_n + \zeta_n$, $\sigma_n = \beta_n + \zeta_n$, where ζ_n is totally monotone.

It is instructive to follow out the construction of χ in a few simple cases.

(i) If $\mu_n = 1$ for all n, then $\lambda_{p,s}$ is 0 for $s < p$ and 1 for $s = p$; χ_q is 0 for $0 \leqslant x < 1$ and 1 for $x = 1$, for every q; and χ is the same function.

(ii) If $\mu_n = (n+1)^{-1}$ then, for $0 \leqslant s \leqslant p$,

$$\lambda_{p,s} = \frac{p!}{s!\,(p-s)!} \Delta^{p-s} \frac{1}{s+1} = \frac{p!}{s!\,(p-s)!} \frac{(p-s)!\,s!}{(p+1)!} = \frac{1}{p+1},$$

and $\qquad \chi_q(0) = 0, \qquad \chi_q(x) = \dfrac{r+1}{q+1} \qquad \left(\dfrac{r}{q} \leqslant x < \dfrac{r+1}{q},\ x > 0\right).$

The limit function χ is x.

(iii) If
$$\mu_n = \binom{n+k}{k}^{-1} = \frac{\Gamma(k+1)\Gamma(n+1)}{\Gamma(n+k+1)},$$
where $k > 0$, then a straightforward calculation gives

$$\binom{q+k}{k} \sum_{s \leqslant qx} \lambda_{q,s} = \sum_{s \leqslant [qx]} \binom{q-s+k-1}{k-1} = \binom{q+k}{k} - \binom{q-[qx]+k-1}{k},$$
$$\sum_{s \leqslant qx} \lambda_{q,s} = 1 - \binom{q+k}{k}^{-1} \binom{q-[qx]+k-1}{k} \to 1-(1-x)^k,$$

and
$$\mu_n = \int x^n\, d\chi = k \int x^n(1-x)^{k-1}\, dx.$$

(iv) If $\mu_n = a^n$, where $0 < a < 1$, then

$$\chi_q(0) = 0, \quad \chi_q(x) = \sum_{0 \leqslant s \leqslant qx} \binom{q}{s} a^s (1-a)^{q-s} \quad (0 < x \leqslant 1).$$

It follows from Theorem 138 that χ is 0 for $0 \leqslant x < a$ and 1 for $a < x \leqslant 1$.

It is sometimes convenient† to modify the definition of $\chi_q(x)$ slightly. We defined $\chi_q(x)$ as a step-function which has a jump $\lambda_{q,r}$ at $x = r/q$. We may eliminate the discontinuities by straight lines connecting the angles of the graph. This process gives a $X_q(x)$ continuous except perhaps for $x = 0$, where it has a jump $\lambda_{q,0}$, and with a derivative $q\lambda_{q,r}$ for $(r-1)/q < x < r/q$; and it is plain that $X_q(x) \to \chi(x)$ when $q \to \infty$ appropriately.

11.10. Inclusion and equivalence of $\mathfrak{H}$ methods. The general problem of the inclusion or 'relative strength' of two $\mathfrak{H}$ methods is difficult, and its solution, which has been effected all but completely by Rogosinski and Fuchs, depends upon the study of the 'Mellin transforms' $M(z) = \int t^z \, d\chi(t)$, associated with the methods, for complex z. The problem is much simplified if the moment constants μ_n and μ'_n of the methods do not vanish for any n; and we confine ourselves to this case, in which the solution can be stated very simply in terms of μ_n and μ'_n.‡

In what follows, then, we assume that

$$(11.10.1) \qquad \mu_n \neq 0, \qquad \mu'_n \neq 0 \quad (n = 0, 1, 2, \ldots).$$

If the transformations are $\lambda = \delta\mu\delta$, $\lambda' = \delta\mu'\delta$, then

$$\delta\frac{1}{\mu}\delta . \lambda = \delta\frac{1}{\mu}\delta . \delta\mu\delta = \delta\frac{1}{\mu}\mu\delta = \delta\delta = \mathrm{I},$$

the identity. Thus

$$\lambda^{-1} = \delta\frac{1}{\mu}\delta, \qquad \lambda' = \delta\mu'\delta, \qquad \lambda'\lambda^{-1} = \delta\mu'\delta . \delta\frac{1}{\mu}\delta = \delta\frac{\mu'}{\mu}\delta,$$

so that $\lambda'\lambda^{-1}$ is the $\mathfrak{H}$ transformation formed from μ'_n/μ_n. But, in order that $(\mathfrak{H}, \mu')$ should include $(\mathfrak{H}, \mu)$, it is necessary and sufficient that $\lambda s \to l$ should imply $\lambda' s = \lambda'\lambda^{-1}(\lambda s) \to l$, i.e. that $\lambda'\lambda^{-1}$ should be regular; and this is so if, and only if, μ'_n/μ_n is a regular moment constant.

THEOREM 209. *Suppose that $(\mathfrak{H}, \mu)$ and $(\mathfrak{H}, \mu')$ are two regular $\mathfrak{H}$ methods subject to* (11.10.1). *Then, for $(\mathfrak{H}, \mu')$ to include $(\mathfrak{H}, \mu)$, it is necessary and sufficient that μ'_n/μ_n should be a regular moment constant. For the two*

† See, for example, § 11.16.

‡ If $(\mathfrak{H}, \mu)$ and $(\mathfrak{H}, \mu')$ both sum a series, then the sums are necessarily the same. For if $t_m = \mathfrak{H}s_n$, $t'_m = \mathfrak{H}'s_n$ are the means of s_n corresponding to the two methods, and $t_m \to s$, $t'_m \to s'$, then $\mathfrak{H}'t_m = \mathfrak{H}t'_m$ by Theorem 197; $\mathfrak{H}'t_m \to s$ and $\mathfrak{H}t'_m \to s'$, because the methods are regular; and therefore $s = s'$. Thus any two regular $\mathfrak{H}$ methods are 'consistent' in the sense of § 4.2.

methods to be equivalent, it is necessary and sufficient that μ_n'/μ_n and μ_n/μ_n' should both be regular moment constants. In particular, for $(\mathfrak{H},\mu)$ to be equivalent to the identity, it is necessary and sufficient that μ_n and $1/\mu_n$ should both be regular moment constants.

In the next section we apply Theorem 209 to some important special cases. We make repeated use of the simple theorem which follows.

THEOREM 210. *Sums, differences, and products of moment constants are themselves moment constants. The product of two regular moment constants is a regular moment constant.*

The assertion about sums and differences is obvious, and we need only consider products. First,

$$\Delta\mu_n\mu_n' = \mu_n\Delta\mu_n'+\Delta\mu_n.\mu_{n+1}',$$

$$\Delta^2\mu_n\mu_n' = \mu_n\Delta^2\mu_n'+2\Delta\mu_n\Delta\mu_{n+1}'+\Delta^2\mu_n.\mu_{n+2}',\ldots,$$

so that $\Delta^p\mu_n \geqslant 0$ and $\Delta^p\mu_n' \geqslant 0$ imply $\Delta^p\mu_n\mu_n' \geqslant 0$. Hence the product of two totally monotone μ_n is totally monotone.

Next, if μ_n and μ_n' are moment constants,

$$\mu_n\mu_n' = \alpha_n\alpha_n'+\beta_n\beta_n'-\alpha_n\beta_n'-\alpha_n'\beta_n,$$

where $\alpha_n,\ldots$ are totally monotone. Hence the product of two moment constants is a moment constant.

Thirdly, $\mu_0 = 1$ and $\mu_0' = 1$ imply $\mu_0\mu_0' = 1$.

Finally, we have to show that $\Delta^m\mu_0 \to 0$ and $\Delta^m\mu_0' \to 0$ imply $\Delta^m\mu_0\mu_0' \to 0$, and it is plainly sufficient to prove this when μ_n and μ_n' are totally monotone. Now

$$\Delta^m\mu_0\mu_0' = \sum_{r=0}^{m}\binom{m}{r}\Delta^r\mu_0\Delta^{m-r}\mu_r' = \sum_{r=0}^{R}+\sum_{r=R+1}^{m} = S_1+S_2,$$

say. We can choose R so that $\Delta^r\mu_0 < \epsilon$ for $r > R$, when

$$S_2 \leqslant \epsilon\sum_{r=0}^{m}\binom{m}{r}\Delta^{m-r}\mu_r' = \epsilon\mu_0',$$

by (11.5.5); and $S_1 \to 0$ when R is fixed and $m \to \infty$. Hence $\Delta^m\mu_0\mu_0' \to 0$; and this completes the proof of the theorem.†

11.11. Mercer's theorem and the equivalence theorem for Hölder and Cesàro means. If λ and λ' are two equivalent $\mathfrak{H}$ methods $(\mathfrak{H},\mu)$ and $(\mathfrak{H},\mu')$, then we write $\lambda \equiv \lambda'$. In particular we write $\lambda \equiv \mathrm{I}$

† Alternatively we might have used Theorem 208.

when λ is equivalent to the identity. Plainly $\lambda^{(1)}\lambda^{(2)}\ldots\lambda^{(r)} \equiv \mathrm{I}$ if $\lambda^{(s)} \equiv \mathrm{I}$ for $s = 1, 2,\ldots, r$.

If

$$\mu_n = \frac{\alpha n+1}{\beta n+1} = \frac{\alpha}{\beta}+\frac{\beta-\alpha}{\beta}\frac{1}{\beta n+1}, \tag{11.11.1}$$

where α and β are positive, then either μ_n or $(\alpha/\beta)-\mu_n$ is totally monotone, $\mu_0 = 1$, and $\Delta^m\mu_0 \to 0$, so that μ_n is a regular moment constant. Since the same is true of μ_n^{-1}, $\lambda \equiv \mathrm{I}$. The transformation

$$t_m = \alpha s_m+(1-\alpha)\frac{s_0+s_1+\ldots+s_m}{m+1}$$

of § 5.9 is the $\mathfrak{H}$ transformation corresponding to

$$\mu_n = \alpha+\frac{1-\alpha}{n+1} = \frac{\alpha n+1}{n+1},$$

and μ_n and μ_n^{-1} are both regular moment constants. We thus obtain another proof of Mercer's Theorem 51.

If $\mu_n \neq 0$ for all n, and μ_n' is a finite product

$$\mu_n' = \mu_n \prod \frac{\alpha_s n+1}{\beta_s n+1} = \mu_n \prod \mu_n^{(s)},$$

where the α and β are all positive, so that $\lambda^{(s)} \equiv \mathrm{I}$, then μ_n'/μ_n and μ_n/μ_n' are both regular moment constants, and $\lambda' \equiv \lambda$. If k is a positive integer, and

$$\mu_n = \frac{1}{(n+1)^k}, \qquad \mu_n' = \binom{n+k}{k}^{-1} = \frac{k!}{(n+1)(n+2)\ldots(n+k)},$$

then

$$\frac{\mu_n'}{\mu_n} = \frac{2n+2}{n+2}\cdot\frac{3n+3}{n+3}\cdots\frac{kn+k}{n+k},$$

and each factor is of the form (11.11.1), so that $\lambda' \equiv \lambda$. This gives another proof of the equivalence theorem (Theorem 49).

We defined Hölder means of non-integral order in § 11.4, and it is natural to ask whether the equivalence theorem can be given a corresponding extension.

THEOREM 211. *The* (C, k) *and* (H, k) *means are equivalent for* $k > -1$.

We have to prove that

$$\rho_n^{(k)} = \binom{n+k}{k}^{-1}(n+1)^k, \qquad \sigma_n^{(k)} = \frac{1}{\rho_n^{(k)}} = (n+1)^{-k}\binom{n+k}{k}$$

are regular moment constants. Since

$$\frac{\rho_n^{(k)}}{\rho_n^{(k+1)}} = \frac{n+k+1}{(k+1)n+k+1},$$

it is sufficient to prove this in some interval $s < k \leqslant s+1$ of values of k, where $s \geqslant -1$ and, since integral values of k are already accounted for, we may suppose that $0 < k < 1$. Now

$$\rho_n^{(k)} = (n+1)^{k-1}\frac{\Gamma(k+1)\Gamma(n+2)}{\Gamma(n+k+1)} = \Gamma(k+1)+(n+1)^{k-1}u_n, \tag{11.11.2}$$

where

$$u_n = \Gamma(k+1)\left\{\frac{\Gamma(n+2)}{\Gamma(n+k+1)}-(n+1)^{1-k}\right\}$$

$$= k(k-1)\int x^n\left\{x(1-x)^{k-2}-\left(\log\frac{1}{x}\right)^{k-2}\right\}dx,$$

so that u_n is a moment constant. Since $(n+1)^{k-1}$ is also a moment constant, it follows from (11.11.2) and Theorem 210 that $\rho_n^{(k)}$ is a moment constant.

Secondly, $\rho_0^{(k)} = 1$. Finally, u_n is the moment constant of an absolutely continuous χ, so that $\chi(+0) = 0$ and $\Delta^m u_0 \to 0$; and $(n+1)^{k-1}$ satisfies the corresponding condition. Hence $\Delta^m \rho_0^{(k)} \to 0$, and $\rho_n^{(k)}$ is a regular moment constant. The proof for $\sigma_n^{(k)}$ is similar.

We may prove in a similar way that the (C, k) and (H, k) methods are equivalent to the $\mathfrak{H}$ methods corresponding to either of

$$\mu_n = \frac{\Gamma(k+a)\Gamma(n+a)}{\Gamma(a)\Gamma(n+k+a)}, \quad \mu_n = \left(\frac{a}{n+a}\right)^k$$

for any positive k and a. Or again we may prove that

$$\binom{n+\alpha+\beta}{\alpha+\beta}\Big/\binom{n+\alpha}{\alpha}\binom{n+\beta}{\beta} = \frac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta+1)}\frac{\Gamma(n+\alpha+\beta+1)\Gamma(n+1)}{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}$$

and its reciprocal are regular moment constants, and so complete the proof of the theorem stated in the note on § 5.8.

It is also interesting to work out the actual expressions of $\rho_n^{(k)}$ and $\sigma_n^{(k)}$ as moment constants. Suppose, for example, that k is integral. Then the formal solution of

$$\rho_n^{(k)} = k!\frac{(n+1)^k}{(n+1)(n+2)\dots(n+k)} = \int_0^1 x^n\,d\chi = \int_0^\infty e^{-nt}\,d\phi \tag{11.11.3}$$

is

$$\phi(x) = \frac{k!}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{(s+1)^k}{(s+1)(s+2)\dots(s+k)}\frac{e^{xs}}{s}\,ds;$$

and the integral may be calculated as a sum of residues. We find that

$$\phi(0) = 0, \quad \phi(+0) = k!, \quad \phi(t) = e^{-t}\left(\frac{d}{dt}\right)^k\{e^t(1-e^{-t})^k\} \quad (t>0), \tag{11.11.4}$$

a form of $\phi(t)$ which may be verified directly.

Similarly we find that $\sigma_n^{(k)}$ is expressible by an integral with

$$\phi(0) = 0, \quad \phi(+0) = \frac{1}{k!}, \quad \phi'(t) = \frac{1}{k!(k-1)!}\left[\left(\frac{d}{dw}\right)^k(w\log w)^{k-1}\right]_{w=e^t}.$$

These formulae remain true for non-integral k if the operation of differentiation is interpreted appropriately. Thus (11.11.4) still holds for $0 < k < 1$ if $k!$ is replaced by $\Gamma(k+1)$ and the derivatives are defined in the manner of Riemann and Liouville.

It should be observed, finally, that we have shown that the Hölder and Cesàro means are equivalent in the range $-1 < k < 0$, when neither is regular.

11.12. Some special cases. (1) For the (H, k) method, with $k > 0$,

$$\mu_n = \frac{1}{(n+1)^k} = \frac{1}{\Gamma(k)} \int x^n \left(\log \frac{1}{x}\right)^{k-1} dx = \int x^n \phi(x)\, dx.$$

Here χ, the integral of ϕ, is absolutely continuous.

(2) For the (C, k) method, with $k > 0$,

$$\mu_n = \frac{\Gamma(k+1)\Gamma(n+1)}{\Gamma(n+k+1)} = k \int x^n (1-x)^{k-1}\, dx;$$

and $\chi(x) = 1-(1-x)^k$ is again absolutely continuous.

(3) These examples suggest that the 'strength' of an $\mathfrak{H}$ method will depend upon the 'smallness' of μ_n, increasing as μ_n becomes smaller. But this principle, though valid up to a point, must not be interpreted strictly, the relations between two moment constants which govern their relative efficiency being of a more subtle character.†

Thus $\mu_n = a^n$, where $0 < a < 1$, corresponds to the method (E, q) with $q = (1-a)/a$, and tends to 0 more rapidly than any $(n+1)^{-k}$; but it is not true that (E, q) includes (C, k), even when q is large and k small. The two methods are in fact 'incomparable'. Suppose, for example, that $k = 1$. Then it is easy to verify that, if $\mu_n = (n+1)^{-1}$ and $\mu'_n = a^n$, neither of $\rho_n = \mu_n/\mu'_n$ and $\sigma'_n = \mu'_n/\mu_n$ is a moment constant. This is obvious for ρ_n, since $\rho_n \to \infty$, and we need only verify it for σ_n.

If $\sigma_n = (n+1)a^n$ were a moment constant then (since a^n is one) we should have

$$na^n = \int x^n\, d\chi = \chi(1) - n \int x^{n-1} \chi(x)\, dx = n \int x^{n-1}\{\chi(1)-\chi(x)\}\, dx$$

for $n > 0$. Dividing by n and replacing n by $n+1$, we obtain

$$a^n = \frac{1}{a} \int x^n \{\chi(1)-\chi(x)\}\, dx = \int x^n\, d\chi_1 \quad (n \geqslant 0),$$

where χ_1 is absolutely continuous. But $a^n = \int x^n\, d\chi_2$, where χ_2 is 0 in $(0, a)$ and 1 in $(a, 1)$, and the dual expression of a^n contradicts Theorem 203. We thus see, as we have proved directly in § 9.8, that there are series summable (C, 1) but not summable (E, q) for any q. The argument is easily adapted to any positive k.

† See the remarks at the beginning of § 11.10, and the notes at the end of the chapter.

(4) If $\mu_n \to 0$ too rapidly, it cannot be a regular moment constant. This is shown by

THEOREM 212. *There is no regular moment constant μ_n such that $c^n\mu_n \to 0$ for every c.*

If μ_n is a moment constant then

$$\mu_{n+2} = \int x^{n+2}\,d\chi = \chi(1)-(n+2)\int x^{n+1}\chi\,dx \tag{11.12.1}$$

$$= \chi(1)-(n+2)\chi_1(1)+(n+2)(n+1)\int x^n\chi_1\,dx = (n+2)(n+1)\int x^n\phi\,dx,$$

where

$$\phi(x) = \chi_1(x)-\chi_1(1)+(1-x)\chi(1), \qquad \chi_1(x) = \int_0^x \chi(t)\,dt, \tag{11.12.2}$$

so that $\phi(x)$ is absolutely continuous. We consider the function

$$f(w) = \int \frac{\phi(x)}{w-x}\,dx$$

of the complex variable $w = u+iv$. We have

$$f(w) = \int \phi(x)\sum \frac{x^n}{w^{n+1}}\,dx = \sum \frac{\mu_{n+2}}{(n+2)(n+1)}\frac{1}{w^{n+1}}$$

for large w, and the series is convergent for all $w \neq 0$. Hence $f(w)$ defines an integral function of $1/w$; in particular it is regular on $0 < w < 1$. But, if $0 < u < 1$, then

$$\frac{1}{2\pi i}\{f(u-iv)-f(u+iv)\} = \frac{1}{\pi}\int \frac{v\phi(x)}{(x-u)^2+v^2}\,dx \to \phi(u) \tag{11.12.3}$$

when $v \to +0$; and so $\phi(u) = 0$ and $\mu_n = 0$ for $n \geqslant 2$.

Also $\phi'(x) = 0$ for $0 < x < 1$, and therefore, by (11.12.2),

$$\chi(x)-\chi(1) = 0$$

for almost all x, so that

$$\mu_1 = \int x\,d\chi = \chi(1)-\int \chi\,dx = 0.$$

Thus the sequence (μ_n) is μ_0, 0, 0,...; and this is not regular, whether $\mu_0 = 0$ or $\mu_0 \neq 0$, since either (11.7.7) or (11.7.6) is violated.

It follows from the last remark that $\mu_n = (n!)^{-1}$ is not a moment constant. It is easily verified that

$$\Delta^p\frac{1}{0!} = \frac{(-1)^p}{p!}\left\{1-\frac{p^2}{1!}+\frac{p^2(p-1)^2}{2!}-\dots\right\} = L_p(1),$$

where $L_p(x)$ is Laguerre's polynomial. It is known that

$$L_p(1) = e^{\frac{1}{2}}\pi^{-\frac{1}{2}}p^{-\frac{1}{4}}\cos(2p^{\frac{1}{2}}-\tfrac{1}{4}\pi)+O(p^{-\frac{3}{4}})$$

for large p, so that $\Delta^r\mu_0$ is not of fixed sign.

It is simpler to prove that a totally monotone μ_n satisfying the condition of Theorem 212 must vanish for $n \geqslant 1$. For if $\mu_n = \int x^n\,d\chi$ and χ is an increasing function not constant in $0 < x \leqslant 1$, then there is an interval (a,b) in $(0,1)$ for which $a > 0$, $\chi(b)-\chi(a) = \omega > 0$; and $\mu_n \geqslant \omega a^n$.

11.13. Logarithmic cases. We now consider a form of μ_n which leads to an $\mathfrak{H}$ method included in and weaker than all the (C,k) methods of positive order. If $a > 1$, $l > 0$, then

$$\frac{1}{\{\log(n+a)\}^l} = \frac{1}{\Gamma(l)}\int\limits_0^\infty t^{l-1}e^{-t\log(n+a)}\,dt = \frac{1}{\Gamma(l)}\int\limits_0^\infty \frac{t^{l-1}}{(n+a)^t}\,dt$$

$$= \frac{1}{\Gamma(l)}\int\limits_0^\infty \frac{t^{l-1}\,dt}{\Gamma(t)}\int\limits_0^\infty u^{t-1}e^{-(n+a)u}\,du = \int\limits_0^\infty e^{-(n+1)u}\psi(u)\,du = \int\limits_0^1 x^n\phi(x)\,dx,$$

where $$\phi(x) = \psi\left(\log\frac{1}{x}\right), \qquad \psi(u) = \frac{e^{(1-a)u}}{u\Gamma(l)}\int\limits_0^\infty \frac{t^{l-1}u^t}{\Gamma(t)}\,dt.$$

The inversions are legitimate because all the functions are positive. It follows that

$$\mu_n = \left\{\frac{\log a}{\log(n+a)}\right\}^l \quad (a > 1,\ l > 0) \tag{11.13.1}$$

is a regular moment constant.

It is not difficult to prove that these methods are weaker than (C,k) or (H,k) for every positive k, but the details of the proof are a little tiresome. Let us suppose for simplicity that $l = 1$, and assume that, as was stated near the end of § 11.11, (H,k) is equivalent to the H method with

$$\mu_n' = \left(\frac{a}{n+a}\right)^k. \tag{11.13.2}$$

We have to show that, if μ_n and μ_n' are defined by (11.13.1), with $l = 1$, and (11.13.2), then μ_n'/μ_n is a regular moment constant and μ_n/μ_n' is not. The second assertion is obvious because $\mu_n/\mu_n' \to \infty$. On the other hand,

$$N^{-k}\log N = -\frac{d}{dk}N^{-k} = \frac{1}{\Gamma(k)}\int\limits_0^\infty e^{-Nt}t^{k-1}\left\{\frac{\Gamma'(k)}{\Gamma(k)} - \log t\right\}\,dt;$$

and it follows from this formula, with $N = n+a$, that $(n+a)^{-k}\log(n+a)$ is a moment constant. Thus μ_n'/μ_n is a moment constant, which is plainly regular.

Since $\mu_n \to 0$, μ_n^{-1} is not a moment constant. It follows that the methods sum some divergent series.

11.14. Exponential cases. It is also interesting to define regular methods $(\mathfrak{H}, \mu)$ stronger than any method (C, k). In such a case μ_n must tend to 0 more rapidly than any power of n; but we have seen in § 11.12 that it must not tend to 0 too rapidly. It is natural, after the examples of §11.12, to consider the case $\mu_n = e^{-An^\alpha}$, where $A > 0$, $0 < \alpha < 1$. We prove

THEOREM 213. *If $A > 0$, $0 < \alpha < 1$, then $\mu_n = e^{-An^\alpha}$ is a regular moment constant corresponding to an increasing χ.*

If $\mu(y) = e^{-Ay^\alpha} = e^{-\nu(y)}$, then $\nu' > 0$, $\nu'' < 0$, $\nu''' > 0$,..., and so

$$\mu' = -e^{-\nu}\nu' < 0, \qquad \mu'' = e^{-\nu}(\nu'^2 - \nu'') > 0,$$
$$\mu''' = -e^{-\nu}(\nu'^3 - 3\nu'\nu'' + \nu''') < 0, ...,$$

so that the successive derivatives of μ alternate in sign. It follows that μ_n is totally monotone, and

(11.14.1) $$\mu_n = \int t^n \, d\chi,$$

where χ increases with t. Also $\chi(1) = \mu_0 = 1$. Hence, in order to prove the theorem, it is only necessary to show that $\Delta^m \mu_0 \to 0$, or, what is equivalent, that $\chi(+0) = \chi(0)$.

Since $\mu^{(1)}(n) = \mu(\frac{1}{2}n)$ is also totally monotone,

$$\mu(\tfrac{1}{2}n) = \mu^{(1)}(n) = \int u^n \, d\chi^{(1)}(u) = \int t^{\frac{1}{2}n} \, d\chi^{(1)}(t^{\frac{1}{2}}),$$

with an increasing $\chi^{(1)}$; so that

(11.14.2) $$\mu(n) = \int t^n \, d\chi^{(1)}(t^{\frac{1}{2}})$$

for $n = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, ...$ We may suppose χ and $\chi^{(1)}$ normalized, and then, comparing (11.14.1) and (11.14.2), and remembering Theorem 203, we see that $\chi^{(1)}(t^{\frac{1}{2}}) = \chi(t)$. Hence (11.14.1) is true whenever n is an integral multiple of $\frac{1}{2}$. Repeating the argument, we see that

(11.14.3) $$\mu(y) = \int t^y \, d\chi(t)$$

whenever y is an integral multiple of $\frac{1}{2}$, of $\frac{1}{4}$, of $\frac{1}{8}$,.... It follows by continuity that it is true for all positive y. Finally,

$$\mu(0) - \mu(y) = \left(\int_0^\delta + \int_\delta^1\right)(1 - t^y)\, d\chi \geqslant (1 - \delta^y)\{\chi(\delta) - \chi(0)\}$$

for $0 < \delta < 1$, and therefore, making $\delta \to 0$,

$$\chi(+0) - \chi(0) \leqslant \mu(0) - \mu(y).$$

Since $\mu(y)$ is continuous, it follows, on making $y \to 0$, that

$$\chi(+0) = \chi(0) = 0;$$

and this completes the proof of the theorem.

The function $\chi(t)$ is absolutely continuous; thus

$$\mu(y) = \int_0^1 t^y \phi(t)\,dt = \int_0^\infty e^{-yu}\psi(u)\,du,$$

and there is no difficulty in finding explicit analytical expressions for $\psi(u)$ and $\phi(t)$. Thus

$$\phi(t) = \frac{A}{2\sqrt{\pi}} t^{-1}\left(\log\frac{1}{t}\right)^{-\frac{3}{2}} e^{-A^2/(4\log 1/t)}$$

when $\alpha = \frac{1}{2}$. When $\alpha = \frac{1}{3}$ we can show, by the use of Liouville's formula,

$$\int_0^\infty\int_0^\infty e^{-u-v-n/(uv)} u^{-\frac{2}{3}} v^{-\frac{1}{3}}\,dudv = \frac{2\pi}{\sqrt{3}} e^{-3n^{\frac{1}{3}}},$$

that
$$\phi(t) = t^{-1}\psi\left(\log\frac{1}{t}\right), \qquad \psi(u) = \frac{A^{\frac{3}{2}}}{3\pi} u^{-\frac{3}{2}} K_{\frac{1}{3}}(2\,.\,3^{-\frac{3}{2}} A^{\frac{3}{2}} u^{-\frac{1}{2}}),$$

where K_ν is the real cylinder function of the third kind. Generally, the inversion formula for Laplace transforms leads to

$$\psi(u) = \sum \frac{(-A)^p}{p!} \frac{u^{-\alpha p-1}}{\Gamma(-\alpha p)} = -\frac{1}{\pi}\sum (-A)^p \frac{\sin\alpha p\pi\,\Gamma(1+\alpha p)}{p!} u^{-\alpha p-1} = u^{-1}W(u^{-\alpha}),$$

where W is an integral function.

We conclude this section by proving

Theorem 214. *The method* $(\mathfrak{H},\mu)$ *of Theorem* 213 *includes all* (C,k) *methods.*

We take $A = 1$, $\alpha = \frac{1}{2}$ to simplify our formulae; the essentials of the proof are not affected. We begin by proving that

$$\rho_n = \left(\frac{n+a}{a}\right)^k e^{-n^{\frac{1}{2}}}$$

is a regular moment constant for any integral k and sufficiently large $a = a(k)$. We write

$$\rho(t) = \left(\frac{t+a}{a}\right)^k e^{-t^{\frac{1}{2}}} = e^{-\nu(t)}, \qquad \nu(t) = t^{\frac{1}{2}} - k\log\frac{t+a}{a}.$$

Then
$$(-1)^{p-1}\nu^{(p)}(t) = \frac{1\,.\,3\ldots(2p-3)}{2^p} t^{-p+\frac{1}{2}} - \frac{k(p-1)!}{(t+a)^p},$$

where $1\,.\,3\ldots(2p-3)$ is to be interpreted as 1 when $p = 1$. The right-hand side will be positive for $p = 1, 2, \ldots$ and all positive t if

$$\frac{(t+a)^{2p}}{t^{2p-1}} > \left\{2k\sqrt{\pi}\frac{\Gamma(p)}{\Gamma(p-\frac{1}{2})}\right\}^2.$$

The minimum of the left-hand side, for varying t, is $(2p)^{2p}(2p-1)^{-2p+1}a > 2pa$, while the right-hand side behaves like $4k^2\pi p$ for large p. Hence $(-1)^{p-1}\nu^{(p)}(t) > 0$ for $p = 1, 2, \ldots$, $t > 0$, when a is sufficiently large. It follows, by the argument used in the proof of Theorem 213, that ρ_n is totally monotone. Also $\rho_0 = 1$; and we can prove, as there, that $\Delta^m\rho_0 \to 0$. Hence the method $(\mathfrak{H},\rho)$ is a regular $\mathfrak{H}$ method when a is sufficiently large.

Finally, as in § 11.11,
$$\sigma_n = \left(\frac{an+a}{n+a}\right)^k$$
defines a method $(\mathfrak{H},\sigma)$ equivalent to the identity. Hence
$$\tau_n = \rho_n \sigma_n = (n+1)^k e^{-n^{\frac{1}{2}}}$$
is the moment constant of a regular $\mathfrak{H}$ method; and hence the method $(\mathfrak{H},\mu)$, with $\mu_n = e^{-n^{\frac{1}{2}}}$, includes all (C, k) methods.

11.15. The Legendre series for $\chi(x)$. We return to the proof of § 11.9, in which $\chi(x)$ is constructed as the limit of a sequence of step-functions $\chi_q(x)$. It is interesting to have other analytical expressions for $\chi(x)$, and one of the most natural is its expansion as a series of Legendre polynomials. We suppose that μ_n is a regular moment constant, so that $\chi(0) = \chi(+0) = 0$ and $\chi(1) = 1$. If we write
$$t = \tfrac{1}{2}(1+x), \qquad \chi(t) = \theta(x),$$
so that $-1 \leqslant x \leqslant 1$ and
$$\mu_n = \int_{-1}^{1} \left(\frac{1+x}{2}\right)^n d\theta,$$
then θ is of bounded variation, and continuous at $x = -1$; and its Legendre series $\sum c_m P_m(x)$ converges to $\tfrac{1}{2}\{\theta(x-0)+\theta(x+0)\}$ for $-1 < x < 1$, to 0 for $x = -1$, and to $\theta(1-0)$ for $x = 1$.

The coefficients c_m are given by
$$c_m = (m+\tfrac{1}{2}) \int_{-1}^{1} \theta(x)P_m(x)\,dx = (m+\tfrac{1}{2})\varpi_m(1)-(m+\tfrac{1}{2}) \int_{-1}^{1} \varpi_m(x)\,d\theta,$$
where
$$\varpi_m(x) = \int_{-1}^{x} P_m(t)\,dt.$$
Thus
$$c_0 = 1-\tfrac{1}{2} \int_{-1}^{1} (1+x)\,d\theta = 1-\int_{0}^{1} t\,d\chi = \mu_0-\mu_1, \tag{11.15.1}$$
$$c_m = -(m+\tfrac{1}{2}) \int_{-1}^{1} \varpi_m(x)\,d\theta = -(m+\tfrac{1}{2}) \int_{0}^{1} \varpi_m(2t-1)\,d\chi \tag{11.15.2}$$
for $m > 0$. Now
$$\varpi_m(2t-1) = \int_{-1}^{2t-1} P_m(u)\,du = 2\int_{0}^{t} P_m(2w-1)\,dw \tag{11.15.3}$$
and
$$P_m(2w-1) = (-1)^m\left\{1-\frac{m+1}{1}\frac{m}{1}w+\frac{(m+1)(m+2)}{1.2}\frac{m(m-1)}{1.2}w^2-\dots\right\}.$$

Substituting in (11.15.1)–(11.15.3), we find

$$\chi(t) = \sum c_m P_m(2t-1),$$

where $c_0 = \mu_0 - \mu_1$ and

$$c_m = 2(-1)^{m-1}(m+\tfrac{1}{2})\left\{\mu_1 - \frac{m+1}{1}\frac{m}{1}\frac{\mu_2}{2} + \frac{(m+1)(m+2)}{1.2}\frac{m(m-1)}{1.2}\frac{\mu_3}{3} - \dots\right\}$$

for $m > 0$.

If χ is absolutely continuous, $\chi' = \phi$, and $\psi(x) = \phi\{\tfrac{1}{2}(1+x)\}$, then the Legendre series of $\psi(x)$ is $\sum a_m P_m(x)$, where

$$\frac{a_m}{2m+1} = (-1)^m \sum_{k=0}^{m} (-1)^k \binom{m+k}{k}\binom{m}{k}\mu_k = (-1)^m \sum_{k=0}^{m} (-1)^k \frac{(m+k)!}{(m-k)!}\frac{\mu_k}{(k!)^2}.$$

11.16. The moment constants of functions of particular classes. It is natural to ask when χ will be a function of some special class; for example, when it will be absolutely continuous, when it will be the integral of a function of the Lebesgue class L^r, and so on. We confine ourselves here to one theorem whose proof is simple.

THEOREM 215. *In order that*

(11.16.1) $$\mu_n = \int x^n \phi(x)\, dx,$$

where $\phi(x)$ is L^r, with $r > 1$, in $(0, 1)$, it is necessary and sufficient that

(11.16.2) $$(p+1)^{r-1} \sum_{s=0}^{p} |\lambda_{p,s}|^r < H^r,$$

where $\lambda_{p,s}$ is defined by (11.3.4), *and H is independent of p.*

We observe that (11.16.2), by Hölder's inequality, implies

$$\sum |\lambda_{p,s}| < H,$$

so that a μ_n satisfying (11.16.2) is certainly a moment constant.

(a) *The condition is necessary.* For

$$\lambda_{p,s} = \binom{p}{s} \int x^s(1-x)^{p-s}\phi(x)\, dx = \int \rho_{p,s}(x)\phi(x)\, dx,$$

where $\rho_{p,s}(x) = \binom{p}{s} x^s(1-x)^{p-s}$, so that $\sum_{s=0}^{p} \rho_{p,s}(x) = 1$ and

$$\int \rho_{p,s}(x)\, dx = \binom{p}{s}\frac{s!(p-s)!}{(p+1)!} = \frac{1}{p+1}.$$

Hence, again by Hölder's inequality,

$$|\lambda_{p,s}|^r \leqslant \left\{\int \rho_{p,s}(x)\,dx\right\}^{r-1} \int \rho_{p,s}(x)|\phi(x)|^r\,dx,$$

$$(p+1)^{r-1}|\lambda_{p,s}|^r \leqslant \int \rho_{p,s}(x)|\phi(x)|^r\,dx,$$

$$(p+1)^{r-1}\sum_{s=0}^{p}|\lambda_{p,s}|^r \leqslant \int |\phi(x)|^r\left\{\sum_{s=0}^{p}\rho_{p,s}(x)\right\}dx = \int |\phi(x)|^r\,dx.$$

(*b*) *The condition is sufficient.* We define $\chi_p(x)$ as in § 11.9, but with the modification indicated at the end of that section.† Then

$$\chi_p'(x) = p\lambda_{p,s} \quad \left(s = 1, 2,..., p;\ \frac{s-1}{p} < x < \frac{s}{p}\right),$$

$$\int |\chi_p'(x)|^r\,dx = p^{r-1}\sum_{s=1}^{p}|\lambda_{p,s}|^r < H^r,$$

by (11.16.2). It follows that there is a subsequence of p, and a function ϕ of L^r, such that $\chi_p' \to \phi$ weakly and

$$\int_0^x \phi(t)\,dt = \lim \int_0^x \chi_p'(t)\,dt = \lim\{\chi_p(x)-\chi_p(+0)\} = \chi(x)-\lim \lambda_{p,0}.$$

But $(p+1)^{r-1}|\lambda_{p,0}|^r < H$, so that $\lambda_{p,0} \to 0$ and

$$\chi(x) = \int_0^x \phi(t)\,dt, \qquad \mu_n = \int_0^1 t^n\phi(t)\,dt.$$

The proof works, with the appropriate modifications, in the limiting case $r = \infty$, and gives $(p+1)|\lambda_{p,s}| < H$ as a necessary and sufficient condition that χ should be the integral of a bounded function. There is no similarly simple result for the case $r = 1$.

11.17. An inequality for Hausdorff means. In this section we prove an inequality which includes a considerable number of special inequalities, important in the theory of functions of the class L^r. We suppose that χ increases, $\chi(1) = 1$, and $\chi(0) = \chi(+0) = 0$, so that μ_n is totally monotone and the method $(\mathfrak{H}, \mu)$ is regular; and that

$$t_m = \sum_{n=0}^{m}\binom{m}{n}\mu_{n,m-n}\,s_n \tag{11.17.1}$$

is the Hausdorff mean of a positive sequence (s_n).

THEOREM 216. *If* $s_n \geqslant 0$, $r > 1$, *then*

$$\sum t_m^r < \left(\int x^{-1/r}\,d\chi\right)^r \sum s_n^r = H(r)\sum s_n^r, \tag{11.17.2}$$

unless $s_n = 0$ *for all* n *or the transformation reduces to the identity.*

† Using χ, however, instead of X.

It is naturally supposed that $\sum s_n^r$ and $\int x^{-1/r}\,d\chi$ are finite. The integral is not a Riemann-Stieltjes integral, since $x^{-1/r}$ is not bounded, and some generalization of the definition is required. We may define it either as one of the general 'Lebesgue-Stieltjes' type, or as the limit of a Riemann-Stieltjes integral over $(\epsilon, 1)$. The second point of view is the more elementary, but we adopt the first for the sake of conciseness. The constant $H(r)$ is the best possible, but we shall not prove this here. We write

$$e_m = e_m(x) = \sum_{n=0}^{m} \binom{m}{n} x^n(1-x)^{m-n}s_n = \sum_{n=0}^{m} \binom{m}{n} x^n y^{m-n} s_n, \tag{11.17.3}$$

where $0 \leqslant x \leqslant 1$ and $y = 1-x$. Then, by Hölder's inequality,

$$e_m^r \leqslant \sum \binom{m}{n} x^n y^{m-n} s_n^r \left\{\sum \binom{m}{n} x^n y^{m-n}\right\}^{r-1} = \sum \binom{m}{n} x^n y^{m-n} s_n^r, \tag{11.17.4}$$

and so

$$\sum_m e_m^r \leqslant \sum_m \sum_{n\leqslant m} \binom{m}{n} x^n y^{m-n} s_n^r \tag{11.17.5}$$

$$= \sum_n x^n s_n^r \sum_{m\geqslant n} \binom{m}{n} y^{m-n} = (1-y)^{-1} \sum_n s_n^r = x^{-1} \sum_n s_n^r$$

for $0 < x \leqslant 1$. Now

$$t_m = \int \left\{\sum_{n=0}^{m} \binom{m}{n} x^n(1-x)^{m-n} s_n\right\} d\chi = \int e_m(x)\,d\chi. \tag{11.17.6}$$

Hence, by (11.17.5)–(11.17.6) and a form of Minkowski's inequality,

$$\left(\sum t_m^r\right)^{1/r} \leqslant \int \left(\sum e_m^r\right)^{1/r} d\chi \leqslant \left\{H(r) \sum s_n^r\right\}^{1/r}. \tag{11.17.7}$$

This is (11.17.2), but with '$\leqslant$' for '$<$'.

There is inequality at the first stage of (11.17.7) unless

$$e_m(x) = K_m \phi(x),$$

except in a set S of x in which the variation of χ is 0. We must distinguish the cases in which the complementary set S' includes (*a*) an infinite number, (*b*) only a finite number of points. In case (*a*) $e_m(x) = K_m\phi(x)$ for all m and an infinity of x; in case (*b*), χ is a step-function.

(*a*) In this case we write $e_m(x)$ in the form

$$e_m(x) = (1-x+xE)^m s_0 = (1-x\Delta)^m s_0 = \sum_{k=0}^{m} (-1)^k \binom{m}{k} x^k \Delta^k s_0.$$

If s_l is the first s_n which is not 0, then $e_m(x) = 0$ for $m < l$, and

$$e_l(x) = (-1)^l x^l \Delta^l s_0 = s_l x^l,$$

$$e_{l+1}(x) = (-1)^l (l+1) x^l \Delta^l s_0 + (-1)^{l+1} x^{l+1} \Delta^{l+1} s_0, \dots .$$

Now $e_l(x) = K_l \phi(x)$, $e_{l+1}(x) = K_{l+1}\phi(x)$ in S', so that $K_l \neq 0$, $K_{l+1} \neq 0$; and the polynomials $K_{l+1} e_l(x)$ and $K_l e_{l+1}(x)$ are equal for an infinity of values of x. It follows that $\phi(x)$ is a multiple of x^l, and $\Delta^m s_0 = 0$ for $m > l$. Hence s_n is a polynomial in n, which must be zero because $\sum s_n^r$ is convergent.

(*b*) If χ is a step-function with jumps α_k at $x = x_k$, then $x_k \neq 0$ (since the method is regular) and $t_m = \sum \alpha_k e_m(x_k)$. Also

$$(11.17.8) \quad \Big(\sum_m t_m^r\Big)^{1/r} \leqslant \sum_k \alpha_k \Big\{\sum_m e_m^r(x_k)\Big\}^{1/r}$$

$$\leqslant \sum_k x_k^{-1/r} \alpha_k \Big(\sum_n s_n^r\Big)^{1/r} = \big(\sum s_n^r\big)^{1/r} \int x^{-1/r}\, d\chi,$$

by (11.17.7) and (11.17.5). There is inequality in (11.17.4) unless either all the s_n have the same value c, or x is 0 or 1. Since $x_k \neq 0$, and $\sum s_n^r$ is convergent, it follows that there will be inequality in (11.17.8) unless either all the s_n are 0 or all the x_k are 1. But in the second case $\chi(x)$ is 0 for $0 \leqslant x < 1$ and 1 for $x = 1$, and the transformation is the identity.

Examples. (1) If $\chi = 0$ for $0 \leqslant x < a < 1$, $\chi = 1$ for $a \leqslant x \leqslant 1$, then $t_m = e_m(a)$, and is the Euler mean of s_n of order $q = (1-a)/a$. Thus if t_m is the (E, q) mean of s_n, and not all s_n are 0, then

$$\sum t_m^r < (q+1) \sum s_n^r.$$

This is equivalent to (11.17.5), with inequality.

(2) If we take $\chi = t$, we obtain

$$\sum \left(\frac{s_0+s_1+\dots+s_n}{n+1}\right)^r < \left(\frac{r}{r-1}\right)^r \sum s_n^r.$$

More generally, if we take $\chi = 1-(1-t)^k$, where $k > 0$, then t_m is the (C, k) mean of s_n, and

$$\sum t_m^r < \left\{\frac{\Gamma(1+k)\Gamma(1-1/r)}{\Gamma(1+k-1/r)}\right\}^r \sum s_n^r.$$

11.18. Continuous transformations. There are transformations of functions of a continuous variable analogous to the $\mathfrak{H}$ transformations discussed in the preceding sections. We are led to them naturally as follows. Our regular $\mathfrak{H}$ transformations of s_n were defined by

$$(11.18.1) \qquad \Delta^n t_0 = \mu_n \Delta^n s_0,$$

where μ_n is a regular moment constant. Suppose now that $f(x)$ is a function of x regular along the positive real axis, and so expressible in the form

$$f(x) = \sum \frac{f^{(n)}(0)}{n!} x^n = \sum a_n x^n$$

for small x. Then the natural analogue of (11.18.1) is

(11.18.2) $$g^{(n)}(0) = \mu_n f^{(n)}(0),$$

in which case

$$g(x) = \sum \mu_n a_n x^n = \int \sum a_n (xt)^n \, d\chi = \int f(xt) \, d\chi,$$

at any rate for small x.

We are thus led to consider the transformation

(11.18.3) $$g(x) = \int f(xt) \, d\chi(t)$$

(dismissing the considerations which led us to the formula). We shall suppose that $f(x)$ is continuous in any finite $(0, X)$: we shall be interested only in the behaviour of $f(x)$ and $g(x)$ when $x \to \infty$.

If $\chi(t) = 1-(1-t)^k$, where $k > 0$, then

$$g(x) = k \int_0^1 f(xt)(1-t)^{k-1} \, dt = \frac{k}{x^k} \int_0^x (x-u)^{k-1} f(u) \, du$$

is the (C, k) mean of $f(x)$ in the sense of § 5.14. If $\chi(t) = 0$ for $0 \leqslant t < a < 1$, $\chi(t) = 1$ for $a \leqslant t \leqslant 1$, then $g(x) = f(ax)$. This is the analogue of the Euler transformation; and, unlike the corresponding transformation of s_n, it is trivial, since $f(x) \to l$ and $g(x) \to l$ are equivalent.

We prove one theorem only. We suppose, as we may, that $\chi(0) = 0$.

THEOREM 217. *In order that the transformation* (11.18.3) *should be regular, i.e. that $f(x) \to l$ should imply $g(x) \to l$, it is necessary and sufficient that $\chi(1) = 1$ and $\chi(+0) = \chi(0) = 0$.*

If $f(x) = 1$ for all x, then $g(x) = \int d\chi = \chi(1)$. Hence $\chi(1) = 1$ is a necessary condition.

If $f(x) = 1$ for $0 \leqslant x \leqslant \delta$, where $\delta > 0$, and $f(x) = 0$ for $x > \delta$, then $f(x) \to 0$. Also

$$g(x) = \int_0^{\delta/x} d\chi = \chi\left(\frac{\delta}{x}\right) - \chi(0).$$

If the transformation is regular, $g(x) \to 0$, and therefore $\chi(\delta/x) \to \chi(0)$, i.e. $\chi(+0) = \chi(0) = 0$.

It remains to prove the conditions sufficient. Since $f = l$ gives $g = l$ it is enough to prove that $f \to 0$ implies $g \to 0$. Further, since χ is the difference of two bounded increasing functions continuous at 0, it is enough to prove this for an increasing χ. Then, if we choose X so that $|f| < \epsilon$ for $x \geqslant X$, and denote by $M(X)$ the upper bound of $|f|$ for $x \leqslant X$, we have

$$|g(x)| \leqslant \int_0^1 |f(xt)|\, d\chi \leqslant \int_0^{X/x} |f(xt)|\, d\chi + \epsilon \int_{X/x}^1 d\chi$$

$$\leqslant M(X)\{\chi(X/x) - \chi(0)\} + \epsilon \int d\chi < 2\epsilon \int d\chi = 2\epsilon$$

for sufficiently large x.

There is an inequality for $g(x)$, when χ is monotone and $f \geqslant 0$, similar to (11.17.2), viz.

$$\int_0^\infty g^r(x)\, dx < \left(\int_0^1 x^{-1/r}\, d\chi\right)^r \int_0^\infty f^r(x)\, dx. \tag{11.18.4}$$

The proof is similar to that of § 11.17, but rather simpler, owing to the triviality of the Euler transformation. In the particular case $\chi = t$, the inequality becomes

$$\int_0^\infty \left\{\frac{1}{x}\int_0^x f(u)\, du\right\}^r dx < \left(\frac{r}{r-1}\right)^r \int_0^\infty f^r(x)\, dx. \tag{11.18.5}$$

11.19. Quasi-Hausdorff transformations. The theory of $\mathfrak{H}$ transformations depends upon the properties of the transformation δ of § 11.1. There is another transformation of very similar form which also generates interesting transformations. This is the transformation δ^* with matrix

$$|\delta^*| \qquad \begin{pmatrix} 1 & 1 & 1 & . & . & . \\ 0 & -1 & -2 & . & . & . \\ 0 & 0 & 1 & . & . & . \\ . & . & . & . & . & . \end{pmatrix},$$

obtained by exchanging rows and columns in $|\delta|$.

THEOREM 218. *δ^* is its own reciprocal: if $t = \delta^* s$ then $s = \delta^* t$.*

One preliminary remark is wanted. The theorem asserts that, if

$$t_m = (-1)^m \sum_{n=m}^{\infty} \binom{n}{m} s_n \tag{11.19.1}$$

$$= (-1)^m \left\{s_m + (m+1)s_{m+1} + \frac{(m+1)(m+2)}{1.2} s_{m+2} + \ldots\right\},$$

then s_m is expressible similarly in terms of t_n. The series in (11.19.1) and the reciprocal equation are infinite and need not converge. We

can, however, avoid considerations of convergence in this theorem by supposing that $s_n = 0$ for $n > N$, in which case $t_m = 0$ for $m > N$; and Theorem 218 is to be interpreted in this sense.

The proof is similar to the proof of Theorem 196. We have

$$u_m = (-1)^m \sum_{n=m}^{\infty} \binom{n}{m} t_n = (-1)^m \sum_{n=m}^{\infty} (-1)^n \binom{n}{m} \sum_{p=n}^{\infty} \binom{p}{n} s_p \tag{11.19.2}$$

$$= (-1)^m \sum_{p=m}^{\infty} s_p \sum_{n=m}^{p} (-1)^n \binom{n}{m}\binom{p}{n}$$

$$= \sum_{p=m}^{\infty} \binom{p}{m} s_p \sum_{n=m}^{p} (-1)^{n-m} \binom{p-m}{n-m};$$

and the inner sum is 1 if $p = m$ and 0 otherwise, so that $u_m = s_m$.

The convergence of the series (11.19.1) does not necessarily imply that of the reciprocal series. Thus $s_n = a^n$, where $0 < a < 1$, gives $t_m = (-a)^m(1-a)^{-m-1}$, and the reciprocal series does not converge unless $a < \frac{1}{2}$. The double series in (11.19.2) is convergent if

$$\sum_{p=m}^{\infty} \binom{p}{m} |s_p| \sum_{n=m}^{p} \binom{p-m}{n-m} = \sum_{p=m}^{\infty} \binom{p}{m} 2^{p-m} |s_p| < \infty,$$

which is true, for example, when $s_n = O(a^n)$ and $a < \frac{1}{2}$.

We now define the transformation $(\mathfrak{H}^*, \mu)$ by $\lambda^* = \delta^* \mu \delta^*$, where μ, as in § 11.3, is a diagonal transformation. We find, formally,

$$t_m = (-1)^m \sum_{n=m}^{\infty} (-1)^n \binom{n}{m} \mu_n \sum_{p=n}^{\infty} \binom{p}{n} s_p = (-1)^m \sum_{p=m}^{\infty} s_p \sum_{n=m}^{p} (-1)^n \binom{n}{m}\binom{p}{n} \mu_n$$

$$= (-1)^m \sum_{p=m}^{\infty} \binom{p}{m} s_p \sum_{n=m}^{p} (-1)^n \binom{p-m}{n-m} \mu_n = \sum_{p=m}^{\infty} \binom{p}{m} \Delta^{p-m} \mu_m \,.\, s_p.$$

Thus, replacing p by n, the transformation is defined by

$$t_m = \sum \lambda^*_{m,n} s_n, \tag{11.19.3}$$

$$\lambda^*_{m,n} = 0 \quad (n < m), \qquad \lambda^*_{m,n} = \binom{n}{m} \Delta^{n-m} \mu_m \quad (n \geqslant m). \tag{11.19.4}$$

11.20. Regularity of a quasi-Hausdorff transformation. We have now to consider when the transformation $(\mathfrak{H}^*, \mu)$ is regular. We suppose that μ_n is a moment constant.

Suppose first that μ_n is totally monotone, so that χ increases with t. Then $\lambda^*_{m,n} \geqslant 0$ and

$$\sum_n \lambda^*_{m,n} = \sum_{n>m} \binom{n}{m} \int t^m (1-t)^{n-m}\, d\chi = \int \frac{t^m}{\{1-(1-t)\}^{m+1}}\, d\chi = \int \frac{d\chi}{t};$$

and (11.5.2) and (11.5.4) will be satisfied if and only if this integral converges and has the value 1. Condition (11.5.3) is satisfied automatically, since $\lambda^*_{m,n} = 0$ when $m > n$. Thus

(11.20.1) $$\mu_{-1} = \int \frac{d\chi}{t} = 1$$

is a necessary and sufficient condition for regularity.

In the general case it is plain that

$$\sum_n |\lambda^*_{m,n}| \leqslant \int \frac{|d\chi|}{t}.$$

On the other hand, whenever $\sum\limits_n |\lambda^*_{m,n}|$ is bounded, the function

$$\mathrm{X}(t) = \int\limits_0^t \frac{d\chi(u)}{u}$$

(suitably modified at its discontinuities) may be obtained as the limit of

$$\mathrm{X}_q(t) = \sum_{n \geqslant q/t} \lambda^*_{q,n},$$

where q tends to infinity through an appropriate sequence q_i, and

$$W(t) = \int\limits_0^t \frac{|d\chi(u)|}{u}$$

is then obtained from $|\lambda^*_{m,n}|$ as $\mathrm{X}(t)$ is from $\lambda^*_{m,n}$. Hence we obtain

THEOREM 219. *If μ_n is a regular moment constant corresponding to χ, then the conditions*

(11.20.2) $$\int \frac{|d\chi|}{t} < \infty, \qquad \int \frac{d\chi}{t} = 1$$

are necessary and sufficient for the regularity of $(\mathfrak{H}^*, \mu)$.

11.21. Examples. (1) If $\chi(t)$ is 0 for $0 \leqslant t < a < 1$ and a for $a \leqslant t \leqslant 1$, then (11.20.1) is plainly satisfied, and

$$\mu_n = a^{n+1}, \qquad \Delta^{n-m}\mu_m = a^{m+1}(1-a)^{n-m}.$$

In this case

$$t_m = a^{m+1} \sum_{n \geqslant m} \binom{n}{m} (1-a)^{n-m} s_n$$

$$= a^{m+1}\left\{s_m + (m+1)(1-a)s_{m+1} + \frac{(m+1)(m+2)}{2!}(1-a)^2 s_{m+2} + \ldots\right\}.$$

We are thus led to the 'circle' method (γ, a) of § 9.11.

(2) If $\chi = t$, $\mu_n = (n+1)^{-1}$, then

$$\binom{n}{m}\Delta^{n-m}\mu_m = \binom{n}{m}\frac{(n-m)!}{(m+1)(m+2)\dots(n+1)} = \frac{1}{n+1}.$$

The transformation is

$$t_m = \frac{s_m}{m+1}+\frac{s_{m+1}}{m+2}+\frac{s_{m+2}}{m+3}+\dots,$$

which is plainly not regular. The integral (11.20.1) diverges.

(3) If $\chi = lt^{l+1}/(l+1)$, where $l > 0$, then

$$\mu_n = l\int t^{n+l}\,dt = \frac{l}{n+l+1};$$

(11.20.1) is satisfied, and the transformation is regular. In this case

$$t_m = l\left\{\frac{s_m}{m+l+1}+\frac{(m+1)s_{m+1}}{(m+l+1)(m+l+2)}+\right.$$
$$\left.+\frac{(m+1)(m+2)s_{m+2}}{(m+l+1)(m+l+2)(m+l+3)}+\dots\right\}. \tag{11.21.1}$$

In particular $l = 1$ gives

$$t_m = (m+1)\left\{\frac{s_m}{(m+1)(m+2)}+\frac{s_{m+1}}{(m+2)(m+3)}+\dots\right\}. \tag{11.21.2}$$

It may be shown that the transformations corresponding to different positive l are all equivalent, and that each is equivalent to $(C,1)$. There are transformations similarly related to (C,k) for any $k > 0$.

NOTES ON CHAPTER XI

§§ 11.1–3. The class of transformations $\lambda = \delta\mu\delta$ was first studied by Hurwitz and Silverman, *TAMS*, 18 (1917), 1–20, who identified it with the class of transformations permutable with H. They were concerned primarily with transformations

$$\alpha_0 \mathrm{I}+\alpha_1\mathrm{H}+\alpha_2\mathrm{H}^2+\dots,$$

where I is the identity and $f(z) = \sum \alpha_n z^n$ is an analytic function regular at the origin, and proved that the transformation is regular if $f(z)$ is regular for $|z-\frac{1}{2}| \leqslant \frac{1}{2}$ and $f(1) = 1$. In particular they proved that the (H,k) and (C,k) transformations are regular transformations λ.

Hausdorff [(A), *MZ*, 9 (1921), 74–109] rediscovered the class λ and developed the more complete theory set out here, in which the class is linked with the 'moment problem for a finite interval'. In particular he proved the fundamental Theorem 206. The proofs of this chapter are mostly derived from this paper or a later one [Hausdorff (B)] in *MZ*, 16 (1923), 220–48. An intervening paper in *MZ*, 9 (1921), 280–99, deals with generalizations in different directions. There is a concise account of the theory in Widder, ch. 3.

§ 11.4. Theorem 200 was proved by Hurwitz and Silverman, *l.c. supra*.

§§ 11.6–7. Hausdorff (A). Hausdorff attributes the definition of a totally mono-

tone sequence to Schur. I have arranged the proof of Theorem 201 in accordance with suggestions of Dr. Bosanquet.

§ 11.8. Theorem 203 is a generalization of the familiar theorem that the system $1, x, x^2, \ldots$ is 'complete in $L(0,1)$'.

§ 11.9. Hausdorff (A, B) gave a number of proofs of Theorem 206. The proof here is substantially the first in (B), which is also given, in rather different form, in Widder, 101–4. The arrangement here has been prompted by suggestions of Dr. Aronszajn and Dr. Bosanquet.

There is a proof of Helly's theorem in Widder, 28–9.

Hausdorff's work is closely related to that of S. Bernstein and Widder on totally monotone functions. We may say that $f(x)$ is totally monotone in $(0,\infty)$ if $(-1)^p f^{(p)}(x) \geqslant 0$ for $x > 0$: thus e^{-x} is totally monotone. It was proved by Bernstein that a necessary and sufficient condition for $f(x)$ to be totally monotone is that

$$f(x) = \int_0^\infty e^{-xt}\, d\chi(t),$$

where $\chi(t)$ is increasing and bounded. This is easily deducible from Hausdorff's theorem, but Bernstein's work was independent and his methods different. We can also (though not quite so simply) deduce Hausdorff's theorem from Bernstein's. For fuller information and references see Widder, ch. 4.

§ 11.10. The main results of Rogosinski and Fuchs will be found in Rogosinski, *PCPS*, 38 (1942), 166–92, and Fuchs, *OQJ*, 16 (1945), 64–77. If T and T′ are *regular* $\mathfrak{H}$ methods, then, in order that T′ should include T, it is necessary and sufficient that $T' = \Theta T$, where Θ is a regular $\mathfrak{H}$ method. If T′ and T are *any* $\mathfrak{H}$ methods, T′ includes T, and $\mu_n \neq 0$ except for a sequence (n_k) of n such that $\sum n_k^{-1} < \infty$, then $T' = \Theta T$, where Θ is a regular $\mathfrak{H}$ method. It is not known whether the condition on (n_k) is the best possible, but Fuchs, *PCPS*, 40 (1944), 189–96, has shown that the theorem stated for regular methods is not true for all methods without reservation.

Hille and Tamarkin, *PNAS*, 19 (1933), 573–7, state a considerable number of more special theorems concerning inclusion.

§ 11.11. The equivalence of (H, k) and (C, k), for general $k > -1$, was first proved by Hausdorff (A), 89–90.

There is an accurate discussion of the inversion formulae referred to at the end of the section in Burkill, *PLMS* (2), 25 (1926), 513–24, and Widder, ch. 2. We may be content to calculate $\rho_n^{(k)}$ and $\sigma_n^{(k)}$ formally and verify the results independently. This is easy for integral k, but rather more troublesome for general k.

§ 11.12. For (11.12.3) see Titchmarsh, *Fourier integrals*, 30–1, or Widder, 338–41.

There is a full account of the Laguerre polynomials in Szegö, *Orthogonal polynomials* (New York, 1939), chs. 5 and 8.

§§ 11.13–14. There is a fuller account of these logarithmic and exponential forms of μ_n in Hausdorff (A).

§ 11.15. See Hausdorff (B), 227–31. Hausdorff's point of view is rather different.

For the expansion of $P_n(2w-1)$ in powers of w ('Murphy's formula') see Hobson, *Spherical and ellipsoidal harmonics* (Cambridge, 1931), 22, or Whittaker and Watson, 311–12.

The inverse of the last formula of the section is

$$\mu_m = \frac{a_0}{m+1} + \frac{ma_1}{(m+1)(m+2)} + \frac{m(m-1)a_2}{(m+1)(m+2)(m+3)} + \dots$$

§ 11.16. There are fuller discussions of Theorem 215 and related theorems in Hausdorff (B), 233–40, and Widder, 109–13. For 'weak convergence' (an idea due to F. Riesz) see Littlewood, 45–9.

§ 11.17. Hardy, *JLMS*, 18 (1943), 46–50. Hardy proves there that $H(r)$ is the best possible constant.

There is a full discussion of the Lebesgue-Stieltjes integral in Saks, *Theory of the integral* (ed. 2, Warsaw, 1937), ch. 3. The properties used in this section are stated in *Inequalities*, 152–7. The form of Minkowski's inequality required is Theorem 201 of *Inequalities* (p. 148), restated for Stieltjes integrals in accordance with pp. 155–6.

§ 11.18. For the general theory of continuous Hausdorff transformations see Rogosinski, *PCPS*, 38 (1942), 344–63, and Fuchs and Rogosinski, *OQJ*, 14 (1943), 27–48.

For (11.18.4) we must use Theorem 202 of *Inequalities* (again restated for Stieltjes integrals).

§ 11.20. The integral $\int t^{-1}\,d\chi$ must again be regarded either as a Lebesgue-Stieltjes integral or as the limit of an integral over $(\epsilon, 1)$. We require the theorem that

$$\sum \int a_n(x)\,d\chi = \int \{\sum a_n(x)\}\,d\chi$$

whenever χ is an increasing function, $a_n(x) \geqslant 0$, and either side is finite. This is a very special case of a theorem stated by Widder, 26, and proved by Saks (l.c. under § 11.17, 76–80). Here $a_n(x)$ is continuous for each n, and $\sum a_n(x)$ is uniformly convergent in any interval $(\epsilon, 1)$, and it is easy to prove what is wanted on the basis of the more elementary definition.

§ 11.21. Hardy, *PCPS*, 20 (1921), 304–7, proved that the transformation (11.21.2) is equivalent to (C, 1), but this theorem is practically a special case of one proved earlier by Knopp. For generalizations see the papers of Hardy and Littlewood and of Knopp cited under § 6.7.

XII

WIENER'S TAUBERIAN THEOREMS

12.1. Introduction. In this chapter we return to the 'Tauberian' theorems whose general character we explained in § 7.1. We have already proved a considerable number of such theorems, for example, in §§ 6.1–3, in Ch. VII (which was entirely occupied with them), in §§ 9.6–7, and in § 9.13; but our methods of proof have varied, and the different methods which we have used may seem at first sight to have little connexion with one another. Here we give an account of a general theory, due in the main to Wiener, which enables us to present most of these special theorems as parts of a systematic whole.

We begin by restating Theorem 92, viz.

(*A*) *if* $s_n \to s$ (A) *and* $s_n = O(1)$, *then* $s_n \to s$ (C, 1).

This is a typical Tauberian theorem which provides a suitable opening for our introductory remarks, but it is more convenient to use the integral analogue. The first hypothesis of (*A*) may be stated in any of the equivalent forms

$$\sum a_n r^n \to s, \quad (1-r)\sum s_n r^n \to s, \quad y\sum s_n e^{-ny} \to s, \quad \frac{1}{x}\sum s_n e^{-n/x} \to s,$$

where $r \to 1$, $y \to 0$, $x \to \infty$; and the conclusion is

$$\frac{1}{x}\sum_{n\leqslant x} s_n \to s.$$

The integral analogue is

(*B*) *if*

$$\frac{1}{x}\int e^{-t/x}F(t)\,dt \to l \tag{12.1.1}$$

and $F(t) = O(1)$, *then*

$$\frac{1}{x}\int_0^x F(t)\,dt \to l, \tag{12.1.2}$$

and it is this theorem which we take as our text.

It may be well to interpolate two remarks whose substance is all but obvious after Ch. VII: see in particular § 7.1.

(i) Theorem (*B*) is the 'corrected converse' of the 'Abelian' theorem

(*B'*) (12.1.2) *implies* (12.1.1),

without any additional hypothesis on $F(t)$. This theorem is simple: for if, as we may, we take $l = 0$, then

$$F_1(t) = \int_0^t F(u)\,du = o(t)$$

implies $\displaystyle \int e^{-t/x}F(t)\,dt = \frac{1}{x}\int e^{-t/x}F_1(t)\,dt = o\Big(\frac{1}{x}\int e^{-t/x}t\,dt\Big) = o(x).$

(ii) Theorem (A) is a trivial corollary of Theorem (B). For, if we assume (B), and take $F(t) = s_n$ for $n \leqslant t < n+1$, then

$$\frac{1}{x}\int e^{-t/x}F(t)\,dt = \frac{1}{x}\sum s_n \int_n^{n+1} e^{-t/x}\,dt = (1-e^{-1/x})\sum s_n e^{-n/x},$$

so that the first hypotheses of (A) and (B) are equivalent. Since the second hypotheses and the conclusions are obviously equivalent, (A) follows from (B). The argument is much the same as one used in § 7.2.

The first hypothesis and the conclusion of (B) are each of the form

(12.1.3) $$P_G(F)\ :\ \frac{1}{x}\int G\Big(\frac{t}{x}\Big)F(t)\,dt \to l\int G(t)\,dt.$$

In the hypothesis

$$G(t) = G_1(t) = e^{-t}, \qquad \int G_1(t)\,dt = 1,$$

and in the conclusion

$$G(t) = G_2(t) = 1\ \ (0 \leqslant t \leqslant 1), \quad 0\ \ (t > 1), \qquad \int G_2(t)\,dt = 1.$$

Thus (B) may be stated in the form

(12.1.4) $$P_{G_1}(F)\ .\ F(t) = O(1) \to P_{G_2}(F);$$

and it seems likely that any theorem of this form will have important Tauberian consequences.

If we make the transformations

$$t = e^\tau, \qquad x = e^\xi, \qquad F(e^\tau) = f(\tau), \qquad e^\tau G(e^\tau) = g(-\tau),$$

then

$$\frac{1}{x}\int_0^\infty G\Big(\frac{t}{x}\Big)F(t)\,dt = \int_{-\infty}^{\infty} e^{\tau-\xi}G(e^{\tau-\xi})F(e^\tau)\,d\tau = \int_{-\infty}^{\infty} g(\xi-\tau)f(\tau)\,d\tau,$$

$$\int_0^\infty G(t)\,dt = \int_{-\infty}^{\infty} g(-\tau)\,d\tau = \int_{-\infty}^{\infty} g(\tau)\,d\tau.$$

Thus, replacing ξ and τ by x and t, (12.1.3) becomes

(12.1.5) $$P_g(f) : \int g(x-t)f(t)\,dt \to l\int g(t)\,dt,$$

where the range is now $(-\infty,\infty)$; and (12.1.4) becomes

(12.1.6) $$P_{g_1}(f)\ .\ f(t) = O(1) \to P_{g_2}(f),$$

with appropriate g_1 and g_2.† We are thus led to ask for general conditions on g_1 and g_2 under which (12.1.6), or a similar theorem with some

† Actually with

$$g_1(t) = e^{-t}e^{-e^{-t}}; \qquad g_2(t) = e^{-t}\quad (t \geqslant 0), \quad 0 \quad (t < 0).$$

alternative condition on f, may be true. Wiener's fundamental discovery is that there is a special condition on g_1, which we shall call $W(g_1)$, and whose form we shall consider in a moment, which almost enables us to dispense with conditions on g_2. His 'key theorem', in any of its forms, is a theorem of the type

(12.1.7) $$W(g_1)\ .\ P_{g_1}(f)\ .\ R(f) \rightarrow P_{g_2}(f),$$

where $R(f)$ is a 'Tauberian' condition on f, and the conclusion is true, not for a special g_2, or for g_2 restricted like g_1, but for 'all reasonable' g_2. It is no longer in general true that

$$Pg_2(f) \rightarrow Pg_1(f),$$

so that the propositions contained in (12.1.7) for different choices of g_1, g_2 are not all corrected converses of Abelian theorems, though they are still 'Tauberian' in a wider sense. We have already met a result of this character in Theorem 147.

12.2. Wiener's condition. The form of $W(g_1)$ is suggested by the theory of Fourier transforms of functions of the class $L(-\infty,\infty)$.†

It is plain that $P_g(f)$ implies $P_h(f)$ when

$$h(t) = \sum_{m=1}^{n} r_m\, g(t-a_m),$$

and this suggests that the inference may be extended, with proper precautions, to $h(t)$ of the form

(12.2.1) $$h(t) = \frac{1}{\sqrt{(2\pi)}} \int r(u)g(t-u)\, du.‡$$

We are thus led to ask whether, given a g of L, an arbitrary h of L can be expressed in the form (12.2.1), with a kernel r also of L.

We define the Fourier transform of a function $r(t)$ of $L(-\infty,\infty)$ by

(12.2.2) $$R(t) = \frac{1}{\sqrt{(2\pi)}} \int r(u)e^{itu}\, du$$

(and similarly with other letters for r and R). It is familiar§ that, if g and r are L, and h is defined by (12.2.1), then h is also L, and $H(t) = R(t)G(t)$. Thus if g and h are given, and we wish to express h in the form (12.2.1), we are led to define $R(t)$ by

(12.2.3) $$R(t) = \frac{H(t)}{G(t)}$$

† Not by Plancherel's more symmetrical theory for functions of $L^2(-\infty,\infty)$.

‡ Here, and to § 12.7 inclusive, the limits, when not shown, are $-\infty$ and ∞.

§ We state the theorems about Fourier transforms which we need more formally in § 12.3.

and $r(t)$, in some sense, by the reciprocal Fourier formula

(12.2.4) $$r(t) = \frac{1}{\sqrt{(2\pi)}} \int R(t)e^{-itu}\,du.$$

It seems essential, if this solution is to be successful, that

(12.2.5) $$G(t) = \frac{1}{\sqrt{(2\pi)}} \int g(u)e^{itu}\,du \neq 0,$$

i.e. that the Fourier transform of $g(t)$ must not vanish.

We define W as the class of functions which (i) belong to $L(-\infty,\infty)$ and (ii) whose Fourier transforms do not vanish for any t, and we suppose that g_1 belongs to W. We shall then find (a) that any h of L can be expressed in the form (12.2.1), with an r of L, and (b) that $P_g(f)$ implies $P_h(f)$ for any bounded f. That the class W should intervene is, as we have seen, quite natural. What is surprising is that so simple a hypothesis as 'g is W' should be sufficient for so general a conclusion: one might naturally expect any theorem of this kind to be encumbered with more complex conditions on g, particularly in regard to the behaviour of G at infinity. (But see Corrigenda, p. 386.)

Our main object now is to prove Wiener's 'key theorem' in the form

THEOREM 220. *If* (i) g *is* W, (ii) h *is* L, *and* (iii) f *is bounded, then* $P_g(f)$ *implies* $P_h(f)$, *i.e.*

(12.2.6) $$\int g(x-t)f(t)\,dt \to l \int g(t)\,dt$$

implies

(12.2.7) $$\int h(x-t)f(t)\,dt \to l \int h(t)\,dt.$$

We shall deduce Theorem 220 from a theorem of Pitt, viz.

THEOREM 221. *If* (i) g *is* W, (ii) f *is bounded and slowly oscillating, or real, bounded, and slowly decreasing, and* (iii) $P_g(f)$, *i.e.* (12.2.6), *is true, then* $f(x) \to l$ *when* $x \to \infty$.

Here we use the terms 'slowly oscillating' and 'slowly decreasing', not as they were used in § 6.2, but in the alternative sense, appropriate to the interval $(-\infty,\infty)$, referred to in the note on § 6.2: the connexion between the two senses will become clear in § 12.8. We say now that $f(x)$ is 'slowly oscillating' if

(12.2.8) $$f(y)-f(x) \to 0$$

when

(12.2.9) $$y > x, \qquad x \to \infty, \qquad y-x \to 0,$$

and 'slowly decreasing' if it is real and

(12.2.10) $$\underline{\lim}\{f(y)-f(x)\} \geqslant 0$$

under the same conditions. Thus $f(x)$ is slowly oscillating if $f'(x) = O(1)$, and slowly decreasing if $f'(x) > -H$. It will plainly be sufficient to prove Theorem 221 for real and slowly decreasing f.

12.3. Lemmas concerning Fourier transforms. We shall use the following theorems concerning Fourier transforms of functions of L: the proofs of the first three will be found in any of the books on the subject. We write '$G \sim g$' for 'G is the transform of g'.

THEOREM 222. *If g is L, and $G \sim g$, then G is continuous and bounded.*

THEOREM 223. *If g is L and $G \sim g$, then*

$$g(t) = \frac{1}{\sqrt{(2\pi)}} \int G(u)e^{-itu}\, du \quad (\mathrm{C}, 1),$$

i.e. $$g(t) = \lim_{U\to\infty} \frac{1}{\sqrt{(2\pi)}} \int_{-U}^{U} \left(1-\frac{|u|}{U}\right) G(u)e^{-itu}\, du,$$

for almost all t. In particular, if G is null, then g is null.

Actually we need this inversion formula only for the sake of the corollary in the last sentence.

THEOREM 224. *If g and r are L, then*

$$h(t) = \frac{1}{\sqrt{(2\pi)}} \int r(u)g(t-u)\, du$$

is L, and $$H(t) = R(t)G(t).$$

THEOREM 225. *If $P \sim p$, $Q \sim q$, then*

(12.3.1) $$P(t-c) \sim e^{-ict}p(t),$$

(12.3.2) $$P(t-c)Q(t) \sim \frac{e^{-ict}}{\sqrt{(2\pi)}} \int p(t-u)q(u)e^{icu}\, du,$$

and

(12.3.3)
$$P(t-c)\{Q(t)-Q(c)\} \sim \frac{1}{\sqrt{(2\pi)}} \int \{p(t-u)-p(t)\}q(u)e^{ic(u-t)}\, du.$$

Of the last three formulae, (12.3.1) is obvious, and (12.3.2) follows from (12.3.1) and Theorem 224. As regards (12.3.3), we have

$$P(t-c)Q(c) = \frac{P(t-c)}{\sqrt{(2\pi)}} \int q(u)e^{icu}\, du \sim \frac{1}{\sqrt{(2\pi)}} \int p(t)q(u)e^{ic(u-t)}\, du.$$

by (12.3.1), and (12.3.3) follows from this and (12.3.2).

12.4. Lemmas concerning the class U. We denote the class of transforms of functions of L by U; and if

$$G(t) = \frac{1}{\sqrt{(2\pi)}} \int g(u)e^{itu}\, du,$$

where g is L, then we write

$$U(G) = \frac{1}{\sqrt{(2\pi)}} \int |g(u)|\, du.$$

Plainly $|G(t)| \leqslant U(G)$ for all real t.

THEOREM 226. *If G_1 and G_2 are U, then G_1+G_2 and $G_1 G_2$ are U, and*

$$U(G_1+G_2) \leqslant U(G_1)+U(G_2), \qquad U(G_1 G_2) \leqslant U(G_1)U(G_2).$$

This is obvious for G_1+G_2. As regards $G_1 G_2$, it is the transform of

$$h(t) = \frac{1}{\sqrt{(2\pi)}} \int g_1(t-u)g_2(u)\, du$$

by Theorem 224; and

$$U(G_1 G_2) = \frac{1}{\sqrt{(2\pi)}} \int |h(t)|\, dt \leqslant \frac{1}{2\pi} \int dt \int |g_1(t-u)||g_2(u)|\, du$$

$$= \frac{1}{\sqrt{(2\pi)}} \int |g_2(u)|\, du \frac{1}{\sqrt{(2\pi)}} \int |g_1(t-u)|\, du = U(G_1)U(G_2).$$

THEOREM 227. *If G_1 and G_2 are U, and $U(G_2) = d < 1$, then $H = G_1/(1+G_2)$ is U, and*

$$U(H) \leqslant \frac{U(G_1)}{1-d}.$$

Since $|G_2| \leqslant U(G_2) = d < 1$

$$H = \sum (-1)^n G_1 G_2^n = \sum (-1)^n H_n,$$

say, and H_n is U, by Theorem 226. If H_n is the transform of h_n, then

$$\sum_M^N (-1)^n H_n = \frac{1}{\sqrt{(2\pi)}} \int \sum_M^N (-1)^n h_n(u) e^{itu}\, du.$$

Now $$U(H_n) = \frac{1}{\sqrt{(2\pi)}} \int |h_n|\, du \leqslant U(G_1)\{U(G_2)\}^n = d^n U(G_1),$$

by Theorem 226, and

$$\frac{1}{\sqrt{(2\pi)}} \int \left| \sum_M^N (-1)^n h_n \right| du \leqslant \sum_M^N U(H_n) \leqslant U(G_1) \sum_M^N d^n \to 0$$

when M and N tend to infinity. It follows (by the theory of 'strong convergence' of functions of L) that there is an h of L such that

$$\frac{1}{\sqrt{(2\pi)}} \int \left| h - \sum_0^N (-1)^n h_n \right| du \to 0,$$

and that $\{1/\sqrt{(2\pi)}\}\int he^{itu}\,du$ is

$$\lim_{N\to\infty}\frac{1}{\sqrt{(2\pi)}}\int\sum_0^N(-1)^n h_n e^{itu}\,du = \lim_{N\to\infty}\sum_0^N(-1)^n G_1 G_2^n = \frac{G_1}{1+G_2} = H.$$

Thus H is U. Finally,

$$U(H) = \lim_{N\to\infty}\frac{1}{\sqrt{(2\pi)}}\int\left|\sum_0^N(-1)^n h_n\right|du \leqslant U(G_1)\sum d^n = \frac{U(G_1)}{1-d}.$$

Our last lemma concerns special functions which will be important in the proof of Theorem 221. We define $p(t)$ by

$$p(t) = 1-|t| \quad (|t|\leqslant 1), \quad 0 \quad (|t|>1).$$

Then

$$P(t) = \frac{1}{\sqrt{(2\pi)}}\int_{-1}^{1}(1-|u|)e^{itu}\,du = \sqrt{\left(\frac{2}{\pi}\right)}\frac{1-\cos t}{t^2} = \frac{1}{\sqrt{(2\pi)}}\left(\frac{\sin\frac{1}{2}t}{\frac{1}{2}t}\right)^2.$$

Both $p(t)$ and $P(t)$ are L, and they are transforms of one another, so that each is U. We shall also write

(12.4.1) $$k_\lambda(t) = p\left(\frac{t}{2\lambda}\right), \qquad K_\lambda(t) = 2\lambda P(2\lambda t) = \sqrt{\left(\frac{2}{\pi}\right)}\frac{\sin^2\lambda t}{\lambda t^2}$$

for every positive λ. Then K_λ and k_λ are transforms of one another and both functions are U.

THEOREM 228. *If $q_\epsilon(t)$ is defined by*

(12.4.2) $$q_\epsilon(t) = 1 \quad (|t|\leqslant\epsilon), \quad 2-\frac{|t|}{\epsilon} \quad (\epsilon\leqslant|t|\leqslant 2\epsilon), \quad 0 \quad (|t|\geqslant 2\epsilon),$$

then the transform of $q_\epsilon(t)$ is

(12.4.3) $$Q_\epsilon(t) = \sqrt{\left(\frac{2}{\pi}\right)}\frac{\cos\epsilon t-\cos 2\epsilon t}{\epsilon t^2}.$$

Also
$$\frac{1}{\sqrt{(2\pi)}}\int|Q_\epsilon(t)|\,dt \leqslant 3$$

(so that q_ϵ is U), and

$$\int|Q_\epsilon(t-y)-Q_\epsilon(t)|\,dt \to 0$$

when y is fixed and $\epsilon\to 0$.

Here $q_\epsilon(t)$ is the 'trapezoidal' function indicated in Fig. 3, and $q_\epsilon(t) = 2k_\epsilon(t)-k_{\frac{1}{2}\epsilon}(t)$, from which (12.4.3) follows immediately. Next,

$$\frac{1}{\sqrt{(2\pi)}}\int|Q_\epsilon(t)|\,dt \leqslant \frac{1}{\pi}\left(\int\frac{1-\cos 2\epsilon t}{\epsilon t^2}\,dt+\int\frac{1-\cos\epsilon t}{\epsilon t^2}\,dt\right)$$
$$= \frac{3}{\pi}\int\frac{1-\cos u}{u^2}\,du = 3.$$

Finally, if
$$C(t) = \sqrt{\left(\frac{2}{\pi}\right)}\frac{\cos t - \cos 2t}{t^2},$$
then $Q_\epsilon(t) = \epsilon C(\epsilon t)$, C is L, and
$$\int |Q_\epsilon(t-y) - Q_\epsilon(t)|\,dt = \epsilon \int |C(\epsilon t - \epsilon y) - C(\epsilon t)|\,dt = \int |C(u-\epsilon y) - C(u)|\,du,$$
which tends to 0 with ϵ.

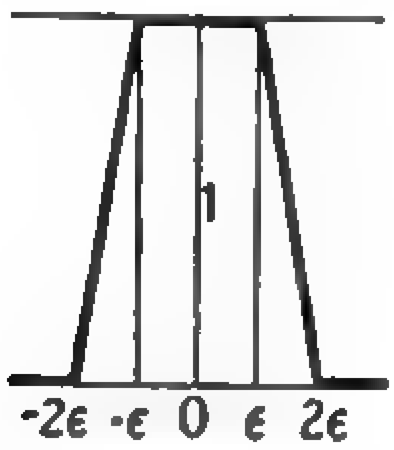

FIG. 3.

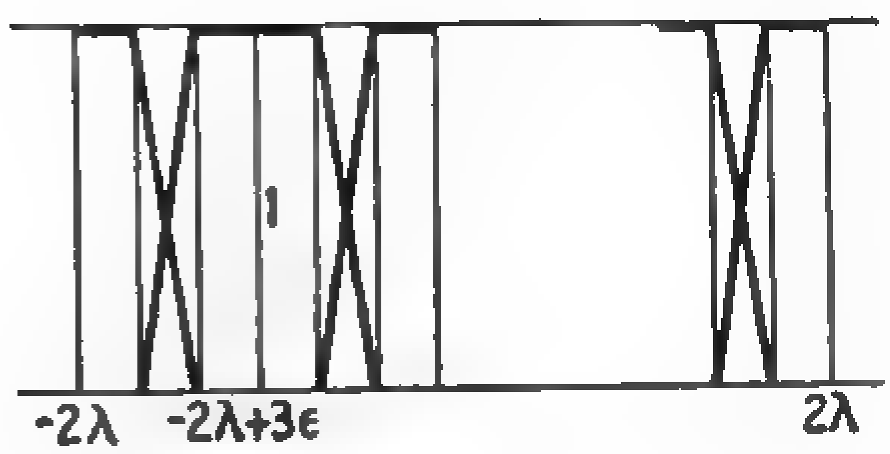

FIG. 4.

12.5. Final lemmas. Our last two lemmas (the first of which is a theorem of importance in itself) contain the kernel of the proofs of Theorems 221 and 220.

THEOREM 229. *If g is W, h is L, and $H(t) = 0$ for $|t| > 2\lambda$, then $H(t)/G(t)$ is U. In particular this is true when $H(t) = k_\lambda(t)$.*

That is to say, if G and H are transforms of functions of L, G does not vanish for any t, and H vanishes for all t outside a finite interval, then H/G is the transform of a function of L.

We divide $(-2\lambda, 2\lambda)$ into N equal intervals by the points
$$t_0 = -2\lambda, \quad t_1 = -2\lambda+3\epsilon, \quad \ldots, \quad t_n = -2\lambda+3n\epsilon, \quad \ldots, \quad t_N = 2\lambda,$$
ϵ being chosen so that $3N\epsilon = 4\lambda$. Then
$$\sum_0^N q_\epsilon(t-t_n) = 1$$
for $|t| \leqslant 2\lambda$ (as is apparent from Fig. 4, in which the vertical lines are at equal distances ϵ). Thus

$$\frac{H(t)}{G(t)} = \sum_{n=0}^{N} \frac{q_\epsilon(t-t_n)H(t)}{G(t)} = \sum_{n=0}^{N} \chi_n(t), \tag{12.5.1}$$

say, for $|t| \leqslant 2\lambda$. Since $G(t) \neq 0$, and $H(t) = 0$ for $|t| > 2\lambda$, this equation holds for all t, and it is sufficient, after Theorem 226, to prove that χ_n is U for each n. We shall prove that this is so for sufficiently small ϵ.

We have $$G(t) = G(t_n)+\{G(t)-G(t_n)\}q_{2\epsilon}(t-t_n)$$
if $|t-t_n| \leqslant 2\epsilon$, since then $q_{2\epsilon}(t-t_n) = 1$; and
$$q_\epsilon(t-t_n)H(t) = 0$$
if $|t-t_n| > 2\epsilon$. Hence
$$\chi_n(t) = \frac{q_\epsilon(t-t_n)H(t)}{G(t_n)+\{G(t)-G(t_n)\}q_{2\epsilon}(t-t_n)} = \frac{G_1\,G_2}{1+G_3},$$
where
$$G_1(t) = \frac{q_\epsilon(t-t_n)}{G(t_n)}, \qquad G_2(t) = H(t), \qquad G_3(t) = \frac{G(t)-G(t_n)}{G(t_n)}q_{2\epsilon}(t-t_n).$$
Here G_1, G_2, and G_3 are U, by Theorems 226 and 228; and it is sufficient, after Theorem 227, to prove that

$$U(G_3) = \frac{1}{|G(t_n)|}U[\{G(t)-G(t_n)\}q_{2\epsilon}(t-t_n)] < 1 \tag{12.5.2}$$

for sufficiently small ϵ.

Now $$\gamma(t) = \{G(t)-G(t_n)\}q_{2\epsilon}(t-t_n)$$
is the Fourier transform of
$$\frac{1}{\sqrt{(2\pi)}}\int \{Q_{2\epsilon}(t-u)-Q_{2\epsilon}(t)\}g(u)e^{it_n(u-t)}\,du,$$
by Theorem 225 (12.3.3). Hence
$$\begin{aligned}U(\gamma) &= \frac{1}{2\pi}\int dt\left|\int \{Q_{2\epsilon}(t-u)-Q_{2\epsilon}(t)\}g(u)e^{it_n(u-t)}\,du\right| \\ &\leqslant \frac{1}{2\pi}\int dt\int |Q_{2\epsilon}(t-u)-Q_{2\epsilon}(t)||g(u)|\,du \\ &= \frac{1}{2\pi}\int |g(u)|\,du\int |Q_{2\epsilon}(t-u)-Q_{2\epsilon}(t)|\,dt.\end{aligned}$$
The inner integral here is bounded, and tends to 0, for any fixed u, when $\epsilon \to 0$, while g is L. It follows that $U(\gamma) \to 0$ when $\epsilon \to 0$. Also $|G(t)|$ has a positive lower bound μ in $(-2\lambda, 2\lambda)$, since $G(t)$ is continuous and does not vanish, and so $U(G_3) \leqslant \mu^{-1}U(\gamma) \to 0$. It follows that (12.5.2) is true for sufficiently small ϵ, and this completes the proof of the theorem.

THEOREM 230. *If*

$$\frac{1}{\sqrt{(2\pi)}}\int K_\lambda(x-t)f(t)\,dt = \frac{1}{\pi}\int \frac{\sin^2\lambda(x-t)}{\lambda(x-t)^2}f(t)\,dt \to l, \tag{12.5.3}$$

for every positive λ, or for some arbitrarily large λ, when $x \to \infty$, and $f(t)$ is bounded and slowly oscillating, or real, bounded, and slowly decreasing, then $f(x) \to l$.

We may suppose $f(t)$ real and slowly decreasing, and $l=0$. If $f(x) \not\to 0$, then there is a positive δ such that one of $f(x) > \delta$ and $f(x) < -\delta$ is true for arbitrarily large x: let us take, for example, the first hypothesis. Then, since $\underline{\lim}\{f(y)-f(x)\} \geqslant 0$ if $x \to \infty$, $y > x$, and $y-x \to 0$, there are an $x_0 = x_0(\delta)$, as large as we please, and an $\eta = \eta(\delta)$, such that

$$f(t) \geqslant \tfrac{1}{2}\delta \quad (x_0 \leqslant x \leqslant t \leqslant x+2\eta).$$

If $\xi = x+\eta$, and M is the upper bound of $|f|$, then

$$\frac{1}{\pi}\int \frac{\sin^2\lambda(\xi-t)}{\lambda(\xi-t)^2} f(t)\,dt \geqslant \frac{\delta}{2\pi}\int\limits_{\xi-\eta}^{\xi+\eta} \frac{\sin^2\lambda(\xi-t)}{\lambda(\xi-t)^2}\,dt - \frac{M}{\pi}\left(\int\limits_{-\infty}^{\xi-\eta} + \int\limits_{\xi+\eta}^{\infty}\right)\frac{\sin^2\lambda(\xi-t)}{\lambda(\xi-t)^2}\,dt$$

$$= \frac{\delta}{\pi}\int\limits_{0}^{\eta} \frac{\sin^2\lambda u}{\lambda u^2}\,du - \frac{2M}{\pi}\int\limits_{\eta}^{\infty} \frac{\sin^2\lambda u}{\lambda u^2}\,du.$$

The first integral here tends to $\frac{1}{2}\pi$, the second to 0, when $\lambda \to \infty$. Hence

$$\frac{1}{\pi}\int \frac{\sin^2\lambda(\xi-t)}{\lambda(\xi-t)^2} f(t)\,dt > \tfrac{1}{4}\delta$$

for sufficiently large λ and arbitrarily large ξ, and

$$\varlimsup_{x\to\infty} \frac{1}{\pi}\int \frac{\sin^2\lambda(x-t)}{\lambda(x-t)^2} f(t)\,dt > 0,$$

in contradiction to (12.5.3). We can prove similarly that the hypothesis $f(x) < -\delta$ leads to a contradiction, and the theorem follows.

It is essential that the hypothesis (12.5.3) should be satisfied for arbitrarily large λ. The Fourier transform $k_\lambda(t)$ of $K_\lambda(t)$ vanishes for $|t| > 2\lambda$, so that

$$\int K_\lambda(x-t)e^{ict}\,dt = 0$$

for $c > 2\lambda$. Thus (12.5.3) is true (with $l=0$) when $f(t) = e^{ict}$ and $c > 2\lambda$; and f is slowly oscillating but does not tend to 0.

12.6. Proof of Theorems 221 and 220. It is now easy to prove Theorem 221. We may suppose $f(t)$ real and slowly decreasing, and $l=0$.

By Theorem 229, k_λ/G is U, i.e.

$$\frac{k_\lambda(t)}{G(t)} \sim r_\lambda(t),$$

where r_λ is L. Hence, by Theorem 224,

$$k_\lambda(t) = \frac{k_\lambda(t)}{G(t)}G(t) \sim \frac{1}{\sqrt{(2\pi)}}\int g(t-u)r_\lambda(u)\,du$$

and, by Theorem 223, for almost all x,

$$K_\lambda(x) = \frac{1}{\sqrt{(2\pi)}}\int g(x-u)r_\lambda(u)\,du.$$

Hence

$$\rho(x) = \int K_\lambda(x-t)f(t)\,dt = \frac{1}{\sqrt{(2\pi)}}\int f(t)\,dt \int g(x-t-u)r_\lambda(u)\,du$$

$$= \frac{1}{\sqrt{(2\pi)}}\int r_\lambda(u)\,du \int g(x-t-u)f(t)\,dt,$$

the inversion being justified by Fubini's theorem. The inner integral here is bounded (since g is L and f is bounded), and tends to 0 (by hypothesis), for each u, when $x \to \infty$; and r_λ is L. Hence $\rho(x) \to 0$ for every λ, and $f(x) \to 0$, by Theorem 230.

It is also easy to deduce Theorem 220: we may suppose $l = 0$. We are given that g is W, that h is L, and that f is bounded. If

$$m(x) = \int h(x-t)f(t)\,dt,$$

then plainly m is bounded. Also

$$m(y)-m(x) = \int \{h(y-t)-h(x-t)\}f(t)\,dt,$$

$$|m(y)-m(x)| \leqslant M\int |h(y-t)-h(x-t)|\,dt = M\int |h(y-x-u)-h(-u)|\,du,$$

where M is again the upper bound of $|f|$. It follows that $m(y)-m(x) \to 0$ when $x \to \infty$, $y-x \to 0$, i.e. that $m(x)$ is slowly oscillating.

Also

$$\int g(x-t)m(t)\,dt$$

$$= \int g(x-t)\,dt \int h(t-u)f(u)\,du = \int f(u)\,du \int g(x-t)h(t-u)\,dt$$

$$= \int f(u)\,du \int g(x-u-w)h(w)\,dw = \int h(w)\,dw \int f(u)g(x-u-w)\,du$$

(the inversions being again justified by Fubini's theorem). The inner integral is bounded, since g is L and f is bounded; and, by hypothesis, it tends to 0, for each w, when $x \to \infty$. Also h is L. Hence

$$\int g(x-t)m(t)\,dt \to 0.$$

But $m(x)$ is bounded and slowly oscillating, and, therefore, by Theorem 221, $m(x) \to 0$, which is (12.2.7).

We have deduced Theorem 220 from Theorem 221. It is also easy to deduce Theorem 221 from Theorem 220, but it will be more convenient to prove this in § 12.8, when we have put the theorems in forms appropriate to the interval $(0, \infty)$.

The condition that $G(t) \neq 0$ is in a sense a *necessary* condition in both theorems. If $G(c) = 0$, $H(c) \neq 0$, and $f(t) = e^{-ict}$, then

$$\int g(x-t)f(t)\,dt = \int g(u)f(x-u)\,du = e^{-icx}\int g(u)e^{icu}\,du = 0$$

for every x, but
$$\int h(x-t)f(t)\,dt = e^{-icx}H(c)$$

does not tend to 0. Thus the condition is necessary in Theorem 220. Since e^{-ict} is also slowly oscillating, the same choice of g and f shows that the condition is necessary in Theorem 221.

12.7. Wiener's second theorem. There is a second theorem of Wiener, concerning Stieltjes integrals, which is also deducible from Theorem 221. We shall make less use of this theorem than of Theorem 220, but it is important theoretically because it can be applied directly to infinite series. We must begin by defining a new class of functions included in and narrower than L.

We shall say that $g(t)$ is M if it is continuous, and

$$\sum \max_{n \leqslant t \leqslant n+1} |g(t)| < \infty \tag{12.7.1}$$

(the sum running from $-\infty$ to ∞). It is plain that any g of M is L, and that (12.7.1) is equivalent to

$$\sum \max_{an+b \leqslant t \leqslant a(n+1)+b} |g(t)| < \infty \tag{12.7.2}$$

for any fixed a (not 0) and b. If g is M, and its transform G does not vanish for any t, then we say that g is W^*: W^* is a subclass of W.

Finally, we consider Stieltjes integrals of the type $\int \phi(t)\,d\alpha(t)$, where $\alpha(t)$ is of bounded variation in any finite interval of t, and

$$\int_t^{t+1} |d\alpha(u)| < H. \tag{12.7.3}$$

THEOREM 231. *If* (i) g *is* W^*, (ii) h *is* M, (iii) α *satisfies* (12.7.3), *and*

$$\int g(x-t)\,d\alpha(t) \to l\int g(t)\,dt, \tag{12.7.4}$$

then

$$\int h(x-t)\,d\alpha(t) \to l\int h(t)\,dt. \tag{12.7.5}$$

We suppose again that $l = 0$, and now write

$$m(x) = \int h(x-t)\,d\alpha(t).$$

Since

$$\int |h(x-t)||d\alpha(t)| = \int |h(u)||d\alpha(x-u)| = \sum \int_n^{n+1} |h(u)||d\alpha(x-u)|$$
$$\leqslant \sum \max_{n \leqslant u \leqslant n+1} |h(u)| \int_n^{n+1} |d\alpha(x-u)| \leqslant H \sum \max_{n \leqslant u \leqslant n+1} |h(u)|,$$

$m(x)$ is bounded. Next,

$$m(y)-m(x) = \int \{h(y-t)-h(x-t)\}\,d\alpha(t) = \int \{h(y-x+t)-h(t)\}\,d\alpha(x-t),$$

$$|m(y)-m(x)| \leqslant \sum \int_n^{n+1} |h(y-x+t)-h(t)||d\alpha(x-t)|$$

$$\leqslant H \sum \max_{n\leqslant t\leqslant n+1} |h(y-x+t)-h(t)|.$$

The last series converges uniformly for $|y-x| \leqslant 1$, since

$$\sum \max_{n\leqslant t\leqslant n+1} |h(t+\tau)|$$

converges uniformly for $|\tau| \leqslant 1$; and each term tends to 0 when $x \to \infty$, $y > x$, and $y-x \to 0$, since $h(t)$ is continuous. Hence $m(y)-m(x) \to 0$ under these conditions, and $m(x)$ is slowly oscillating. Finally,

$$\int g(x-t)m(t)\,dt = \int g(x-t)\,dt \int h(t-u)\,d\alpha(u) = \int d\alpha(u) \int g(x-t)h(t-u)\,dt,$$

the inversion being justified by the convergence of

$$\int |g(x-t)|\,dt \int |h(t-u)||d\alpha(u)|;$$

and this is

$$\int d\alpha(u) \int g(x-u-w)h(w)\,dw = \int h(w)\,dw \int g(x-u-w)\,d\alpha(u).$$

The inner integral is bounded and tends to 0, for each w, when $x \to \infty$; and h is M and *a fortiori* L; so that

$$\int g(x-t)m(t)\,dt \to 0.$$

It now follows from Theorem 221 (as in the proof of Theorem 220) that $m(x) \to 0$.

Theorem 231 has the advantage mentioned at the beginning of this section, but we shall use it comparatively little. It will usually be more convenient to bring our problems, by some preliminary transformation, into a form adapted for the application of Theorem 220 or Theorem 221.

12.8. Theorems for the interval $(0,\infty)$. We now modify our fundamental theorems by an exponential transformation. They then become theorems concerning functions defined over $(0,\infty)$, and it is in this form that they are usually most convenient for application.

We put, for a moment,†

$$e^x = \xi, \quad e^t = \tau, \quad f(t) = F(e^t) = F(\tau), \quad g(-t) = e^t G(e^t) = \tau G(\tau),$$

† Abandoning the use of capital letters for Fourier transforms.

so that $$\int_{-\infty}^{\infty} g(t)\,dt,\qquad \int_{-\infty}^{\infty} g(x-t)f(t)\,dt,\qquad \int_{-\infty}^{\infty} g(t)e^{ixt}\,dt$$

become $$\int_{0}^{\infty} G(\tau)\,d\tau,\qquad \frac{1}{\xi}\int_{0}^{\infty} G\left(\frac{\tau}{\xi}\right)F(\tau)\,d\tau,\qquad \int_{0}^{\infty} G(\tau)\tau^{-ix}\,d\tau.$$

We then replace ξ, τ, F, and G by x, t, f, and g, and the result is as follows. The class of functions g of $L(-\infty,\infty)$ becomes the class $L(0,\infty)$. The class W becomes the subclass of $L(0,\infty)$ for which

$$\int_{0}^{\infty} g(t)t^{-ix}\,dt \neq 0 \tag{12.8.1}$$

for any real x: we still call this class W. The class of slowly oscillating (decreasing) functions f becomes the class which is slowly oscillating (decreasing) in the sense of § 6.2, i.e. the class such that

$$\lim\{f(y)-f(x)\} = 0 \qquad [\varliminf\{f(y)-f(x)\} \geqslant 0]$$

when $$x \to \infty,\qquad y > x,\qquad y/x \to 1.$$

Thus Theorems 220 and 221 become

THEOREM 232. *If g is W, h is L, f is bounded, and*

$$\frac{1}{x}\int g\left(\frac{t}{x}\right)f(t)\,dt \to l\int g(t)\,dt, \tag{12.8.2}$$

then

$$\frac{1}{x}\int h\left(\frac{t}{x}\right)f(t)\,dt \to l\int h(t)\,dt. \tag{12.8.3}$$

THEOREM 233. *If g is W, f is bounded and slowly oscillating, or real, bounded, and slowly decreasing, in the sense of* § 6.2, *and* (12.8.2) *is true, then $f(x) \to l$.*

Here the limits are 0 and ∞. Generally, when integrals are written without limits, the limits will be $-\infty$ and ∞ if the integral involves $x-t$ or e^{ixt}, 0 and ∞ if it involves t/x or t^{-ix}.

To obtain the analogue of Theorem 231, we put

$$\int e^{t}\,d\alpha(t) = \mathrm{A}(e^{t}),$$

and then replace A by α, when (12.7.3) becomes

$$\int_{1}^{e^{t}} \frac{|d\alpha(u)|}{u} < H. \tag{12.8.4}$$

The class M becomes the class of continuous G for which

$$\sum \max_{e^n \leqslant t \leqslant e^{n+1}} |tG(t)| < \infty.$$

Hence, using M in this sense and W^* in the corresponding sense, we obtain

THEOREM 234. *If* (i) g *is* W^*, (ii) h *is* M, (iii) α *satisfies* (12.8.4), *and*

(12.8.5) $$\frac{1}{x}\int g\left(\frac{t}{x}\right) d\alpha(t) \to l \int g(t)\, dt,$$

then

(12.8.6) $$\frac{1}{x}\int h\left(\frac{t}{x}\right) d\alpha(t) \to l \int h(t)\, dt.$$

Finally, we observe that if

$$x = \xi^{-1}, \qquad t = \tau^{-1}, \qquad f(t) = F(\tau), \qquad g(t) = \tau^2 G(\tau),$$

then

$$\int g(t)\, dt = \int G(\tau)\, d\tau, \qquad \frac{1}{x}\int g\left(\frac{t}{x}\right) f(t)\, dt = \frac{1}{\xi}\int G\left(\frac{\tau}{\xi}\right) F(\tau)\, d\tau,$$

and that the classes of functions occurring in the theorems are unchanged. Hence

THEOREM 235. *The results of Theorems* 232 *and* 233 *remain true when* x *tends to* 0 *instead of to* ∞ *in hypotheses and conclusion.*

We can now see how (as was stated in § 12.6) Theorem 221 may be deduced from Theorem 220. It is the same thing to deduce Theorem 233 from Theorem 232. If $f(t)$ is bounded, and we take $h(t) = 1$ for $0 \leqslant t \leqslant 1$, $h(t) = 0$ for $t > 1$, then it follows from Theorem 232 that

(12.8.7) $$\frac{1}{x}\int_0^x f(t)\, dt \to l.$$

If also $f(t)$ is slowly decreasing, then it follows from the integral analogue of Theorem 68 of § 6.2 that $f(x) \to l$.

12.9. Some special kernels. The following special choices of $g(t)$ are particularly important: the first is for $(-\infty, \infty)$, the rest for $(0, \infty)$.

(1) $g(t) = e^{-ct^2}$ $(c > 0)$: $\displaystyle\int g(t) e^{ixt}\, dt = \sqrt{\left(\frac{\pi}{c}\right)} e^{-x^2/4c} \neq 0.$

(2) $g(t) = e^{-t}$: $\displaystyle\int g(t) t^{-ix}\, dt = \Gamma(1-ix) \neq 0.$

(3) $g(t) = 1$ $(0 \leqslant t < 1)$, 0 $(t > 1)$: $\displaystyle\int g(t) t^{-ix}\, dt = \frac{1}{1-ix} \neq 0.$

(4) $g(t) = k(1-t)^{k-1}$ $(0 \leqslant t < 1)$, 0 $(t > 1)$; $k > 0$:

$$\int g(t)t^{-ix}\,dt = k\int_0^1 (1-t)^{k-1}t^{-ix}\,dt = \frac{\Gamma(k+1)\Gamma(1-ix)}{\Gamma(k+1-ix)} \neq 0.$$

(5) $g(t) = \left(\dfrac{\sin t}{t}\right)^2$: $\quad \int g(t)t^{-ix}\,dt = i2^{ix}\Gamma(-1-ix)\sinh\tfrac{1}{2}\pi x \neq 0.$†

(6) $g(t) = \dfrac{d}{dt}\left(\dfrac{\sin t}{t}\right)^2$: $\quad \int g(t)t^{-ix}\,dt = ix2^{1+ix}\Gamma(-2-ix)\cosh\tfrac{1}{2}\pi x \neq 0.$†

(7) If
$$g(t) = \frac{d}{dt}\left(\frac{te^{-t}}{1-e^{-t}}\right),$$
then
$$\int g(t)t^{-ix+\delta}\,dt = (ix-\delta)\int \frac{t^{-ix+\delta}}{e^t-1}\,dt = (ix-\delta)\Gamma(1-ix+\delta)\zeta(1-ix+\delta)$$
for $\delta > 0$, $x \neq 0$; and so, making $\delta \to 0$,
$$\int g(t)t^{-ix}\,dt = ix\Gamma(1-ix)\zeta(1-ix)$$
for $x \neq 0$. If $x = 0$ then the value is -1. Thus the assertion that $g(t)$ is W is equivalent to the theorem that $\zeta(1-ix) \neq 0$.

(8) Finally, we suppose $g_0(t) = [t^{-1}]$ and
$$g(t) = 2g_0(t) - ag_0(at) - bg_0(bt),$$
where a and b are positive and $\log a/\log b$ is irrational. The kernel $g_0(t)$ is not L, since $tg_0(t) \to 1$ when $t \to 0$, but
$$g(t) = \frac{2}{t} - \frac{a}{at} - \frac{b}{bt} + O(1) = O(1)$$
for small t, and $g(t) = 0$ for large t, so that $g(t)$ is L. If $\Re s > 0$ then
$$\gamma_0(s) = \int g_0(t)t^s\,dt = \int [t^{-1}]t^s\,dt = \int [u]u^{-s-2}\,du$$
$$= \frac{1}{s+1}\{(1^{-s-1}-2^{-s-1}) + 2(2^{-s-1}-3^{-s-1}) + \ldots\} = \frac{\zeta(1+s)}{1+s},$$
$$\gamma(s) = \int g(t)t^s\,dt = (2-a^{-s}-b^{-s})\frac{\zeta(1+s)}{1+s}.$$
This equation has been proved for $\Re s > 0$, but both sides are regular for $\Re s > -1$, so that it holds for all such s, and in particular for $s = -ix$, where x is real. Finally, $\gamma(0) = \log a + \log b \neq 0$, and
$$\gamma(-ix) = (2 - e^{ix\log a} - e^{ix\log b})\frac{\zeta(1-ix)}{1-ix} \neq 0$$

† In (5) and (6) we must take the limiting values $\frac{1}{2}\pi$ and -1 when $x = 0$.

when $x \neq 0$, because $\log a/\log b$ is irrational and $\zeta(1-ix) \neq 0$. Thus $g(t)$ is W.

It will be seen that, in the last two examples, the assertion that $g(t)$ is W is equivalent to the theorem that $\zeta(s)$ does not vanish on the line $\sigma = \Re s = 1$. This naturally suggests that the corresponding cases of our theorems will prove important in the theory of the distribution of primes.

The functions (1)–(7) are plainly all L, and so (1)–(8) are all W. It is also plain that the function (1) is M. Finally, if g is continuous, and $O(t^{-1-\delta})$, where $\delta > 0$, for large t, then

$$\sum_{-\infty}^{-1} \max_{e^n \leqslant t \leqslant e^{n+1}} |tg(t)| < H \sum_{-\infty}^{-1} e^n < \infty,$$

$$\sum_{0}^{\infty} \max_{e^n \leqslant t \leqslant e^{n+1}} |tg(t)| < H \sum_{0}^{\infty} e^{-\delta n} < \infty.$$

Hence the functions (1), (2), and (5)–(7) are M, while (4) is M if $k > 1$, but not if $k \leqslant 1$, since then it is discontinuous.

12.10. Application of the general theorems to some special kernels. We now apply our theorems to some of the kernels of § 12.9.

(1) If

$$g(t) = e^{-t}, \qquad h(t) = k(1-t)^{k-1} \quad (0 \leqslant t < 1), \qquad h(t) = 0 \quad (t > 1),$$

and $k > 0$, then Theorem 232 gives

$$\frac{1}{x}\int e^{-t/x} f(t)\,dt \to l \;.\; f(t) = O(1) \to \frac{k}{x^k}\int_0^x (x-t)^{k-1} f(t)\,dt \to l,$$

i.e. $$f(t) \to l \text{ (A)} \;.\; f(t) = O(1) \to f(t) \to l \text{ (C}, k).$$

The special case $k = 1$ is Theorem 92*a*. We saw in Ch. VII how we could deduce all the theorems of § 7.5 from Theorem 92, of which some are direct generalizations, and this leads us to ask whether there are corresponding extensions of Wiener's general theorems. In particular, is it possible (at the price, no doubt, of further restrictions on g) to replace the condition that $f(t)$ is bounded by a one-sided condition $f(t) > -H$? We come back to this question in § 12.12.

If we take $g = k(1-t)^{k-1}$ and $h = \kappa(1-t)^{\kappa-1}$, where $0 < \kappa < k$, for $t < 1$, and $g = h = 0$ for $t \geqslant 1$, we obtain

$$f(t) \to l \text{ (C}, k) \;.\; f(t) = O(1) \to f(t) \to l \text{ (C}, \kappa).$$

If $\kappa > k$, the inference is Abelian and no secondary condition on f is wanted. It is easy to deduce that

$$f(x) = O(1)\ \ (\mathrm{C}, k_1)\ .\ f(x) \to l\ \ (\mathrm{C}, k_2) \to f(x) \to l\ \ (\mathrm{C}, k)$$

for $-1 < k_1 < k < k_2$: this is the form assumed by Theorem 70 for functions of a continuous variable.

(2) If we take $g(t) = e^{-t}$ in Theorem 233, we find that *if* $f(t) \to l$ (A), *and f is bounded and slowly decreasing, then* $f \to l$. This is an imperfect theorem (though in no way a trivial one), since, after Theorem 105, the condition of boundedness is unnecessary. This leads us to ask whether (again probably at the expense of some restriction of g) it may be possible to get rid of the condition of boundedness in Theorem 233.

It is worth while to consider in this connexion why a condition of boundedness is unnecessary in Theorem 106 *a*. We obtained this theorem in Ch. VII as the climax of a rather intricate chain of reasoning, and we consider here only the simplest case, in which $f'(t) = O(t^{-1})$. Then the simple argument of § 7.2 proves

$$f = O(1)\ \ (\mathrm{A})\ .\ f' = O(t^{-1}) \to f = O(1);$$

and so 'f bounded' appears at once as a rather trivial consequence of the other hypotheses.

We cannot expect to answer the question so easily for a general g and a less heavily restricted f: but in § 12.13 we shall prove a theorem, due substantially to Vijayaraghavan, which answers it under fairly general conditions.

(3) If we suppose $k > 1$, and take

$$g = e^{-t}, \qquad h = k(1-t)^{k-1}\ \ (t < 1), \qquad h = 0\ \ (t \geqslant 1),$$

in Theorem 234, we obtain

$$\frac{1}{x}\int e^{-t/x}\, d\alpha(t) \to l\ .\ \int\limits_t^{\epsilon t} \frac{|d\alpha(u)|}{u} < H \to \frac{k}{x^k}\int\limits_0^x (x-t)^{k-1}\, d\alpha(t) \to l.$$

We cannot take $k = 1$ because then h, being discontinuous, is not M. If now $\alpha(t) = s_1+s_2+\ldots+s_n$ for $n \leqslant t < n+1$ we obtain

$$s_n \to l\ \ (\mathrm{A})\ .\ \sum_t^{\epsilon t} \frac{|s_n|}{n} < H \to s_n \to l\ \ (\mathrm{C}, k)$$

for $k > 1$. The second hypothesis is satisfied, in particular, if s_n is bounded. We are thus led directly to a theorem about series, a little more general than Theorem 92 in one way, but weaker in that it asserts summability for $k > 1$ instead of for $k = 1$. We can prove summability

(C, 1), or summability (C, k) for any positive k, when s_n is bounded, by combining the theorem with Theorem 70. In this case there is no particular advantage in using Theorem 234 instead of Theorem 232; but for other g the passage from integrals to series may be less immediate.

(4) It is natural, after § 4.17 (where we were concerned with the similar definitions for series), to express the hypotheses

$$(12.10.1)\quad \frac{2}{\pi y}\int \frac{\sin^2 yt}{t^2} f(t)\,dt \to l, \qquad (12.10.2)\quad \int \left(\frac{\sin yt}{yt}\right)^2 f(t)\,dt \to l,$$

when $y \to 0$, by

$$f(t) \to l \ (\mathrm{R}_2), \qquad \int f(t)\,dt = l \ (\mathrm{R}, 2)$$

respectively. If we choose $g(t)$ as in § 12.9(5), we obtain

$$f(t) \to l \ (\mathrm{R}_2) \ . \ f(t) = O(1) \rightarrow f(t) \to l \ (\mathrm{C}, k)$$

for any $k > 0$. The reverse implication, with R_2 and (C, k) interchanged, is also valid: the R_2 and (C, k) methods are equivalent for bounded f. We shall see in § 12.12 that, when $k = 1$, the condition of boundedness may be replaced by the one-sided condition $f > -H$.

Theorem 233 gives

'$f \to l$ (R_2) . *f is bounded and slowly decreasing* $\rightarrow f \to l$':

this also follows by combining what we have just proved with Theorem 68 *a* (§ 6.2). We shall see in § 12.14 that the condition of boundedness may be dropped.

Theorem 234 gives

$$(12.10.3)\qquad s_n \to s \ (\mathrm{R}_2) \ . \ s_n = O(1) \rightarrow s_n \to s \ (\mathrm{C}, k)$$

for $k > 1$. We may get rid of the restriction $k > 1$, as under (3), by use of Theorem 70.

It will be observed that in none of these cases is the full truth revealed immediately by one of the 'key theorems': a supplementary argument, depending on intrinsically simpler theorems, is necessary in each case. And all the results may be proved by other methods. Thus Szász, using Theorem 94, but without appealing to any of Wiener's theorems, proved that

$$\lim_{\theta\to 0}\frac{4}{\pi\theta}\sum \frac{b_n}{n}\sin^2\tfrac{1}{2}n\theta = l \ . \ nb_n > -H \rightarrow \lim_{x\to\infty}\frac{1}{x}\sum_{n\leqslant x} nb_n = l,$$

and Hardy and Rogosinski afterwards proved this and the reverse implication still more simply. If we replace θ by $2y$ and nb_n by s_n, we obtain (12.10.3), with the one-sided generalization.

Again, we shall see in Appendix III that summability (R_2) implies summability (A), without any reservation on the sequence or function considered. We shall then be able to deduce any of these Tauberian theorems for R_2 from its analogue for A.

(5) The hypothesis (12.10.2) is not, as it stands, one of Wiener's type. If, however,

$$F(t) = \int_0^t f(u)\,du = o(t^2)$$

for large t, then partial integration gives

$$\int \left(\frac{\sin yt}{yt}\right)^2 f(t)\,dt = y\int g(yt)F(t)\,dt,$$

where $g(t)$ is, apart from sign, the function of § 12.9(6). We shall see in Appendix III that the convergence of $\sum n^{-2}\sin^2 ny\,a_n$, for all small y, implies that of $\sum n^{-2}a_n$, and *a fortiori* implies $s_n = o(n^2)$. If we take this, and the corresponding theorem for integrals, for granted, then Theorem 232 gives

$$f \to l\ (\mathrm{R}, 2)\ .\ f = O(1) \rightarrow f \to l\ (\mathrm{C}, k),$$

and similar questions about possible generalization present themselves. We shall see in § 12.16 that these are less simple for this $g(t)$ because it is not of constant sign. We shall, however, prove in Appendix III that summability (R, 2) always implies summability (A), so that the Tauberian theorems for (R, 2), like those for R_2, may be deduced from those for A.

We end this section by two general remarks.

(*a*) It will often happen that the result given by Wiener's theorems is one which may be proved more simply by other methods: this is true, for example, of all the theorems of Ch. VII, and of the theorems referred to under (4) above. The merits of Wiener's method lie in its great power and generality, and the light which it throws on the whole subject; not in simplicity.

(*b*) There will be a complex of Tauberian theorems associated with any particular kernel $g(t)$ and connected by relations of varying simplicity. One of the Wiener theorems concerning g will 'hit the map' in a particular place, which will not always be just the place we want. It will usually be possible to pass from one spot on the map to another by comparatively simple arguments; and it is usually easier to do this than to strain for variations of the general theorems, though such variations have often considerable intrinsic interest.

12.11. Applications to the theory of primes. One of the outstanding applications of Wiener's theorems is to the theory of prime numbers. It had long been familiar that the 'prime number theorem'

$$\pi(x) \sim \frac{x}{\log x} \tag{12.11.1}$$

was 'roughly equivalent' to the theorem (first proved by Hadamard and de la Vallée-Poussin) that

$$\zeta(1+i\tau) \neq 0 \tag{12.11.2}$$

for any real τ. Wiener's theorems enable us to present this equivalence in a much sharper form than was possible before. All previous proofs of the prime number theorem took from the theory of $\zeta(s)$ not merely (12.11.2) but some stronger result such as

$$|\zeta(1+i\tau)| > H(\log|\tau|)^{-K}$$

for large $|\tau|$.

It is known that the prime number theorem is equivalent to

$$\psi(x) = \sum_{n \leqslant x} \Lambda(n) \sim x, \tag{12.11.3}$$

and also (though the proof of this is less familiar) to

$$M(x) = \sum_{n \leqslant x} \mu(n) = o(x).\dagger \tag{12.11.4}$$

We shall deduce (12.11.4) from Theorem 233 and (12.11.2). We choose $g(t)$ as in § 12.9(8), and

$$f(t) = t^{-1}M(t). \tag{12.11.5}$$

We saw in § 12.9 that g is W, and it is obvious that f is bounded. Hence, if we can prove (i) that f is slowly oscillating and (ii) that f and g satisfy (12.8.2) with $l = 0$, then (12.11.4) will follow from Theorem 233. Now

$$f(y)-f(x) = \frac{M(y)}{y}-\frac{M(x)}{x} = \frac{M(y)-M(x)}{x}+M(y)\left(\frac{1}{y}-\frac{1}{x}\right)$$

$$= \frac{1}{x}\sum_{x<n\leqslant y}\mu(n)-\frac{y-x}{x}\frac{M(y)}{y} = O\left(\frac{y-x}{x}\right) = o(1)$$

when $x \to \infty$, $y/x \to 1$. Thus f is slowly oscillating. Finally,

$$\int g_0\left(\frac{t}{x}\right)f(t)\,dt = \int_1^x \left[\frac{x}{t}\right]\frac{M(t)}{t}\,dt = \int_1^x \left\{\sum_{m\leqslant x/t} 1 \sum_{n\leqslant t}\mu(n)\right\}\frac{dt}{t} = \sum_{mn\leqslant x}\mu(n)\int_n^{x/m}\frac{dt}{t}$$

$$= \sum_{mn\leqslant x}\mu(n)\log\frac{x}{mn} = \sum_{q\leqslant x}\log\frac{x}{q}\sum_{n|q}\mu(n) = \log x = o(x),$$

† For the definitions of $\Lambda(n)$ and $\mu(n)$ see Hardy and Wright, Chs. 16–17. We shall give the deduction of (12.11.3) from (12.11.4) in Appendix IV.

since the inner sum is 0 unless $q = 1$, and then it is 1. Hence

$$\int g\left(\frac{t}{x}\right)f(t)\,dt = 2\int g_0\left(\frac{t}{x}\right)f(t)\,dt - a\int g_0\left(\frac{at}{x}\right)f(t)\,dt - b\int g_0\left(\frac{bt}{x}\right)f(t)\,dt = o(x),$$

which is (12.8.2) with $l = 0$. We thus obtain (12.11.4) and the prime number theorem.

12.12. One-sided conditions. In this section we show how it may be possible to replace the condition that the $f(t)$ of Theorem 232 is bounded by the more general condition

(12.12.1) $$f(t) > -H.$$

We confine ourselves for simplicity to the most important special case, in which $h(t)$ is the $g(t)$ of § 12.9 (3); and we shall find it necessary to put additional restrictions on our present $g(t)$.

THEOREM 236. *If* (i) *g is W*, (ii) *$g \geqslant 0$ for all t*, (iii) *there are positive numbers c and K such that*

(12.12.2) $$g(t) \geqslant K$$

for $0 \leqslant t \leqslant c$, (iv) *$f$ satisfies* (12.12.1), *and* (v) *f and g satisfy* (12.8.2), *then*

$$f(x) \to l \quad (C, 1).$$

We may suppose (adding H to f) that $f \geqslant 0$, and that $\int g(t)\,dt = 1$. We write

(12.12.3) $$\sigma(x) = \frac{1}{x}\int_0^x f(t)\,dt,$$

so that

(12.12.4) $$\sigma'(x) = \{f(x) - \sigma(x)\}/x$$

for almost all x. Then

$$\frac{1}{x}\int g\left(\frac{t}{x}\right)f(t)\,dt \geqslant \frac{K}{x}\int_0^{cx} f(t)\,dt = Kc\sigma(cx) \geqslant 0.$$

The left-hand side tends to a limit when $x \to \infty$, and is therefore bounded for large x. Thus $\sigma(x)$ is bounded for large x, and therefore for $x \geqslant 1$.

If $\sigma(x) \leqslant \mu$ for $x \geqslant 1$, then, by (12.12.4),

$$\sigma'(x) \geqslant -x^{-1}\sigma(x) \geqslant -\mu x^{-1}$$

for almost all $x \geqslant 1$. Hence $\sigma(x)$ is slowly decreasing. Next,

$$\int g(u)f(xu)\,du \to l$$

and so $$\frac{1}{x}\int_0^x dy \int g(u)f(yu)\,du \to l.$$

But the left-hand side is

$$\frac{1}{x}\int g(u)\,du\int_0^x f(yu)\,dy = \int g(u)\left\{\frac{1}{xu}\int_0^{xu} f(w)\,dw\right\}du$$

$$= \int g(u)\sigma(xu)\,du = \frac{1}{x}\int g\left(\frac{t}{x}\right)\sigma(t)\,dt.$$

Hence the last integral tends to l. Applying Theorem 233, with σ for f, we find that $\sigma(x) \to l$, i.e. $f(x) \to l$ (C, 1).

The conditions are satisfied, for example, if $g(t)$ is e^{-t} or $\left(\dfrac{\sin t}{t}\right)^2$. The first case gives Theorem 94 *a*, while the second gives

THEOREM 237. *If* $f(t) \to l\,(\mathrm{R}_2)$ *and* $f(t) > -H$, *then* $f(t) \to l$ (C, 1).

This, together with the corresponding result with R_2 and (C, 1) interchanged, is the theorem of Wiener referred to in § 12.10 (4), and proved otherwise by Szász and by Hardy and Rogosinski. The conditions of Theorem 236 are not satisfied when $g(t)$ is the kernel of § 12.9 (6), associated with (R, 2) summability, since this $g(t)$ is not of constant sign.

Theorem 236, unlike Theorems 232 or 233, is one with a specialized h, but this is not a serious disadvantage, since it is usually possible to deduce (12.8.3), with whatever h we may require, from the existence of the (C, 1) limit. The theorem is not true for all g of W: *some* additional condition on g is essential. Suppose, for example, that

$$g(t) = 1-2\log\frac{1}{t}\ \ (0<t<1), \quad 0\ \ (t \geqslant 1).$$

Plainly g is L, and

$$\int g(t)t^{-ix}\,dt = \frac{1}{1-ix}-\frac{2}{(1-ix)^2} = -\frac{1+ix}{(1-ix)^2} \neq 0,$$

so that g is W. If $f(t) = t$, then $f \geqslant 0$ and

$$\frac{1}{x}\int g\left(\frac{t}{x}\right)f(t)\,dt = \frac{1}{x}\int_0^x \left(1-2\log\frac{x}{t}\right)t\,dt = x\int_0^1 \left(1-2\log\frac{1}{u}\right)u\,du = 0,$$

but $f(x) \to \infty$ (C, 1).

12.13. Vijayaraghavan's theorem. Our next theorem is of a different kind and does not depend upon the theory of Fourier transforms.

We are led to it by a remark which we made, in § 12.10(2), about the applications of Theorems 232 and 233. These theorems depend on the hypothesis that $f(t)$ is bounded, but this hypothesis is often redundant in the application. It is therefore important to obtain other theorems in which the boundedness of $f(t)$ appears as a conclusion instead of as a hypothesis. Such a theorem should be of the type 'if $g(t)$ is L, and satisfies, say, conditions (α); $f(t)$, or s_n, satisfies conditions (β); and

$$\frac{1}{x}\int g\left(\frac{t}{x}\right)f(t)\,dt \to l\int g(t)\,dt \quad \left[\text{or } \frac{1}{x}\sum g\left(\frac{n}{x}\right)s_n \to l\int g(t)\,dt\right],$$

or at any rate the sum or integral is bounded; then $f(t)$, or s_n, is bounded'. Conditions (β) will not by themselves imply boundedness, and conditions (α) will not include Wiener's. This section and the next will be occupied by the proof of a theorem of this character, due essentially to Vijayaraghavan. It will be convenient to work, as he does, with series rather than integrals, and to generalize his hypotheses a little.

In what follows, then, we shall be concerned with a method of summation defined by

$$\tau(x) = \sum c_n(x)s_n \to s. \tag{12.13.1}$$

We suppose that

$$c_n(x) \geqslant 0, \qquad c_n(x) \to 0 \quad (x \to \infty), \qquad \sum c_n(x) = 1, \tag{12.13.2}$$

so that the method is totally regular.

THEOREM 238. *Suppose that the following conditions are satisfied.*

(i) *$\phi(u)$ is positive and differentiable for $u \geqslant 1$;*

$$\phi \to \infty, \qquad 0 < \phi' < K, \tag{12.13.3}$$

where K is independent of u,

$$\Phi(u) = \int_1^u \frac{dt}{\phi(t)} \tag{12.13.4}$$

(so that $\Phi \to \infty$ with u).

(ii) *The coefficients c_n have, in addition to those already stated, the properties:*

$$\sum_{n=0}^{M} c_n(x) \to 0 \tag{12.13.5}$$

if $\quad M \to \infty, \quad x \to \infty, \quad \Phi(x)-\Phi(M) \to \infty;$

$$\sum_{n=N}^{\infty} c_n(x) \to 0, \tag{12.13.6}$$

$$\sum_{n=N}^{\infty} c_n(x)\{\Phi(n)-\Phi(N)\} \to 0, \tag{12.13.7}$$

if $\quad N \to \infty, \quad x \to \infty, \quad \Phi(N)-\Phi(x) \to \infty.$†

(iii) *If* $s(t) = s_n$ *for* $n \leqslant t < n+1$, *then*

$$\underline{\lim}\{s(t)-s(u)\} \geqslant 0 \tag{12.13.8}$$

when $\quad u \to \infty, \quad t > u, \quad \dfrac{t-u}{\phi(u)} \to 0;$

(iv) $\tau(x) = \sum c_n(x)s_n$ *is bounded.*

Then s_n *is bounded.*

We shall require a lemma.

THEOREM 239. *If* $s(t)$ *satisfies condition* (iii) *and* $\phi(u)$ *condition* (i) *of Theorem* 238, *then there are positive numbers* a *and* b *such that*

$$s(q)-s(p) > -a\int_p^q \frac{du}{\phi(u)}-b = -a\{\Phi(q)-\Phi(p)\}-b \tag{12.13.9}$$

for $q \geqslant p \geqslant 1$.

It follows from condition (iii) that there are a U and δ such that

$$s(t)-s(u) > -1 \tag{12.13.10}$$

if

$$t \geqslant u \geqslant U, \quad \frac{t-u}{\phi(u)} \leqslant \delta. \tag{12.13.11}$$

If, on the other hand, $u < U$, $t \leqslant U+\delta\phi(U)$, then $s(t)-s(u)$ has a lower bound depending only on U. It follows that there are numbers γ and δ such that

$$s(t)-s(u) > -\gamma \tag{12.13.12}$$

for all t and u for which

$$0 < t-u \leqslant \delta\phi(u). \tag{12.13.13}$$

† That is to say,

$$\sum_0^M c_n(x) < \epsilon, \qquad \sum_N^\infty c_n(x) < \epsilon, \qquad \sum_N^\infty c_n(x)\{\Phi(n)-\Phi(N)\} < \epsilon$$

if x, M, N, $\Phi(x)-\Phi(M)$, $\Phi(N)-\Phi(x)$ are all greater than numbers depending on ϵ. In the applications $x = x(H)$, $M = M(H)$, $N = N(H)$ will be functions of a single parameter H which tends to ∞. We shall not use the full force of our conditions on $\phi(u)$ and c_n, and are content to state the theorem in a form sufficiently general for the applications.

We write

(12.13.14) $p_0 = p, \quad p_1 = p_0+\delta\phi(p_0), \quad \ldots, \quad p_{k+1} = p_k+\delta\phi(p_k), \quad \ldots$

and suppose that $p_r \leqslant q < p_{r+1}$. Then

(12.13.15) $$s(q)-s(p) = \sum_{k=0}^{r-1}\{s(p_{k+1})-s(p_k)\}+s(q)-s(p_r) > -(r+1)\gamma,$$

by (12.13.12), and

(12.13.16) $$\delta r = \sum_{k=0}^{r-1} \frac{p_{k+1}-p_k}{\phi(p_k)}.$$

Now

$$\phi(p_{k+1}) = \phi(p_k)+(p_{k+1}-p_k)\phi'(\xi) \quad (p_k < \xi < p_{k+1}),$$

$$\phi(p_{k+1}) = \phi(p_k)\{1+\delta\phi'(\xi)\} \leqslant (1+\delta K)\phi(p_k),$$

(12.13.17)
$$\delta r \leqslant (1+\delta K)\sum_{k=0}^{r-1}\frac{p_{k+1}-p_k}{\phi(p_{k+1})} \leqslant (1+\delta K)\int_{p_0}^{p_r}\frac{du}{\phi(u)} \leqslant (1+\delta K)\int_{p}^{q}\frac{du}{\phi(u)}.$$

It follows from (12.13.15) and (12.13.17) that

$$s(q)-s(p) > -\gamma\left(\frac{1}{\delta}+K\right)\int_{p}^{q}\frac{du}{\phi(u)}-\gamma,$$

which is (12.13.9).

12.14. Proof of Theorem 238. We have to prove s_n bounded. If s_n tended to ∞ or to $-\infty$, then $\tau(x)$ would do the same (since the transformation is totally regular). It is therefore sufficient to prove that the hypotheses

$$(a)\quad \tau(x) = O(1), \qquad (b)\quad \overline{\lim}|s_n| = \infty,$$

$$(c)\quad s_n \text{ does not tend to } \infty \text{ or to } -\infty,$$

lead to a contradiction. We write

(12.14.1) $$\sigma_1(t) = \max_{n\leqslant t} s_n, \qquad \sigma_2(t) = \max_{n\leqslant t}(-s_n).$$

Then it is plain that $\sigma_1(t)$ and $\sigma_2(t)$ increase with t, that one at least of them tends to ∞, and that there are two possibilities: either (α) $\sigma_1(n) \geqslant \sigma_2(n)$ for an infinity of n, or (β) $\sigma_1(n) < \sigma_2(n)$ for all sufficiently large n. We consider these two possibilities in turn, and show that each leads to contradiction.

Case (α). It is plain that in this case $\sigma_1(t) \to \infty$, and that, given any H, there are M for which

(12.14.2) $$s_M = \sigma_1(M) > 2H, \qquad \sigma_1(M) \geqslant \sigma_2(M).$$

We choose the least $M = M(H)$ satisfying these conditions, and then the least $N = N(H) > M$ such that

(12.14.3) $$s_N \leqslant \tfrac{1}{2}s_M:$$

there are certainly such N when H is large, since otherwise s_n would tend to ∞.

Since $$s_N - s_M > -a\{\Phi(N)-\Phi(M)\}-b,$$

$$a\{\Phi(N)-\Phi(M)\} > s_M - s_N - b \geqslant \tfrac{1}{2}s_M - b > H-b,$$

and so $$\Phi(N)-\Phi(M) \to \infty$$

when $H \to \infty$. Hence, if we define x by

(12.14.4) $$\Phi(x) = \tfrac{1}{2}\{\Phi(M)+\Phi(N)\},$$

then

(12.14.5) $$\Phi(x)-\Phi(M) \to \infty, \qquad \Phi(N)-\Phi(x) \to \infty.$$

We write

(12.14.6) $$\tau(x) = \Big(\sum_{n=0}^{M-1} + \sum_{n=M}^{N-1} + \sum_{n=N}^{\infty}\Big) c_n(x)s_n = \tau_1(x)+\tau_2(x)+\tau_3(x),$$

and estimate τ_1, τ_2, and τ_3 in turn, denoting generally by $\delta(H)$ a function of H which tends to 0 when $H \to \infty$. First,

(12.14.7) $$\tau_1(x) \geqslant -\sigma_2(M)\sum_0^{M-1} c_n \geqslant -\sigma_1(M)\sum_0^{M} c_n \geqslant -\delta(H)\sigma_1(M),$$

by (12.14.1), (12.14.2), and (12.13.5). Secondly, since N is the first n after M which satisfies (12.14.3),

(12.14.8)
$$\tau_2(x) > \tfrac{1}{2}\sigma_1(M)\sum_{M+1}^{N-1} c_n = \tfrac{1}{2}\sigma_1(M)\Big(1-\sum_0^{M} c_n - \sum_N^{\infty} c_n\Big) > \{\tfrac{1}{2}-\delta(H)\}\sigma_1(M),$$

by (12.14.2), (12.13.5), and (12.13.6). Thirdly, if $n \geqslant N$,

(12.14.9) $$s_n - s_{N-1} > -a\{\Phi(n)-\Phi(N-1)\}-b,$$

by (12.13.9). Also $s_{N-1} > \tfrac{1}{2}s_M > H > b+1$ for large H, and

$$\Phi(N)-\Phi(N-1) = \int_{N-1}^{N} \frac{du}{\phi(u)} \to 0,$$

so that $a\Phi(N) < a\Phi(N-1)+1$ for large H. It now follows from (12.14.9) that

$$s_n > -a\{\Phi(n)-\Phi(N-1)\}+1 > -a\{\Phi(n)-\Phi(N)\}$$

for large H, and that

(12.14.10) $$\tau_3(x) > -a\sum_N^{\infty} c_n\{\Phi(n)-\Phi(N)\} > -\delta(H)$$

by (12.13.7). Finally, combining (12.14.7), (12.14.8), and (12.14.10), we obtain

$$\tau(x) > -\delta(H)\sigma_1(M)+\{\tfrac{1}{2}-\delta(H)\}\sigma_1(M)-\delta(H),$$

which tends to infinity with H, in contradiction to the hypothesis that $\tau(x)$ is bounded. Thus case (α) leads to a contradiction.

Case (β). In this case $\sigma_2(n) > \sigma_1(n)$ for all large n, and $\sigma_2(n) \to \infty$. We choose the least $N = N(H)$ such that

$$\sigma_2(n) > \sigma_1(n) \quad (n \geqslant N), \qquad s_N = -\sigma_2(N) < -2H; \tag{12.14.11}$$

and then the last $M = M(H) < N$ for which $s_M \geqslant \frac{1}{2}s_N = -\frac{1}{2}\sigma_2(N)$: there are certainly such M when H is large, since otherwise s_n would tend to $-\infty$. Thus

$$s_M \geqslant -\tfrac{1}{2}\sigma_2(N), \qquad s_n < -\tfrac{1}{2}\sigma_2(N) \quad (M < n \leqslant N). \tag{12.14.12}$$

Then

$$s_N - s_M > -a\{\Phi(N)-\Phi(M)\}-b, \tag{12.14.13}$$

by (12.13.9), and $s_N - s_M \leqslant \frac{1}{2}s_N < -H$, so that

$$a\{\Phi(N)-\Phi(M)\} > H-b,$$

and $\Phi(N)-\Phi(M) \to \infty$ when $H \to \infty$. Hence (12.14.5) is still true when x is defined by (12.14.4).

We now write

$$\tau(x) = \Big(\sum_{n=0}^{M} + \sum_{n=M+1}^{N} + \sum_{N+1}^{\infty}\Big)c_n(x)s_n = \tau_1(x)+\tau_2(x)+\tau_3(x), \tag{12.14.14}$$

and estimate τ_1, τ_2, and τ_3. First,

$$\tau_1(x) \leqslant \sigma_1(M)\sum_0^M c_n \leqslant \sigma_1(N)\sum_0^M c_n \leqslant \sigma_2(N)\sum_0^M c_n \leqslant \delta(H)\sigma_2(N), \tag{12.14.15}$$

by (12.14.1), (12.14.11), and (12.13.5). Secondly,

$$\begin{aligned}\tau_2(x) &= \sum_{M+1}^{N} c_n s_n < -\tfrac{1}{2}\sigma_2(N)\sum_{M+1}^{N} c_n \\ &= -\tfrac{1}{2}\sigma_2(N)\Big(1-\sum_0^M c_n - \sum_{N+1}^{\infty} c_n\Big) \leqslant -\{\tfrac{1}{2}-\delta(H)\}\sigma_2(N),\end{aligned} \tag{12.14.16}$$

by (12.14.12), (12.13.5), and (12.13.6). Thirdly,

$$\begin{aligned}\tau_3(x) &= \sum_{N+1}^{\infty} c_n s_n \leqslant \sum_{N+1}^{\infty} c_n \sigma_1(n) \leqslant \sum_{N+1}^{\infty} c_n \sigma_2(n) \\ &= \sum_{N+1}^{\infty} c_n\{\sigma_2(N)+\sigma_2(n)-\sigma_2(N)\}.\end{aligned} \tag{12.14.17}$$

Now

(12.14.18) $\quad -s_n-\sigma_2(N) = -s_n+s_N < a\{\Phi(n)-\Phi(N)\}+b$

for $n > N$, by (12.13.9), and therefore

(12.14.19) $\quad \sigma_2(n)-\sigma_2(N) < a\{\Phi(n)-\Phi(N)\}+b.$†

Hence, after (12.14.17),

(12.14.20)

$$\tau_3(x) \leqslant \sigma_2(N)\sum_{N+1}^{\infty} c_n + a\sum_{N+1}^{\infty} c_n\{\Phi(n)-\Phi(N)\}+b\sum_{N+1}^{\infty} c_n \leqslant \delta(H)\sigma_2(N),$$

by (12.13.6) and (12.13.7). Finally, it follows from (12.14.15), (12.14.16), and (12.14.20) that

$$\tau(x) \leqslant -\{\tfrac{1}{2}-\delta(H)\}\sigma_2(N),$$

and so that $\tau(x) \to -\infty$ when $H \to \infty$. This is again a contradiction, so that case (β) is disposed of and the theorem proved.

Suppose in particular that

$$c_n = (1-e^{-1/x})e^{-n/x}, \qquad \phi(u) = u, \qquad \Phi(u) = \log u.$$

Then $\quad \displaystyle\sum_0^M c_n = (1-e^{-1/x})\sum_0^M e^{-n/x} = 1-e^{-(M+1)/x} < \frac{M+1}{x} \to 0$

if $M/x \to 0$, i.e. $\log x - \log M \to \infty$; and

$$\sum_N^{\infty} c_n = (1-e^{-1/x})\sum_N^{\infty} e^{-n/x} = e^{-N/x} \to 0$$

if $N/x \to \infty$, i.e. $\log N - \log x \to \infty$. Finally,

$$\sum_{n=N}^{\infty} c_n \log\frac{n}{N} = (1-e^{-1/x})\sum_{\nu=0}^{\infty}\log\Bigl(1+\frac{\nu}{N}\Bigr)e^{-(N+\nu)/x}$$

$$< \frac{e^{-N/x}}{Nx}\sum \nu e^{-\nu/x} < \{x(1-e^{-1/x})\}^{-2}\frac{x}{N}e^{-N/x},$$

which tends to 0 if $x \to \infty$, $\log N - \log x \to \infty$.

Thus the conditions (i) and (ii) of Theorem 238 are satisfied, while (iii) asserts that s_n is slowly decreasing in the sense of § 6.2. It follows that *if $s_n = O(1)$ (A), and s_n is slowly decreasing, then $s_n = O(1)$.* We can now deduce from Theorem 233 that *if $s_n \to s$ (A), and s_n is slowly decreasing, then $s_n \to s$*, which is Theorem 106. Thus Theorem 106 is a corollary of Theorems 233 and 238, though (as we saw in § 12.10 (2)) not of Theorem 233 alone.

As a second example, we may take

$$c_n(x) = \frac{1}{x}g\Bigl(\frac{n}{x}\Bigr), \qquad g(t) = \Bigl(\frac{\sin t}{t}\Bigr)^2$$

† $\sigma_2(n) = \max\limits_{\nu\leqslant n}(-s_\nu) = \max\{\sigma_2(N), -s_{N+1}, \ldots, -s_n\}$. If the maximum arose from $\sigma_2(N)$, (12.14.19) would be trivial; if from one of $-s_{N+1}, \ldots, -s_n$, then it follows from (12.14.18).

(with the same ϕ and Φ). Then (12.13.1) is equivalent to $s_n \to s\,(\mathrm{R}_2)$. In this case

$$x\sum_{1}^{M}\frac{\sin^2(n/x)}{n^2} = O\left(x\sum_{1}^{M}\frac{1}{x^2}\right) = O\left(\frac{M}{x}\right) \to 0,$$

$$x\sum_{N}^{\infty}\frac{\sin^2(n/x)}{n^2} = O\left(x\sum_{N}^{\infty}\frac{1}{n^2}\right) = O\left(\frac{x}{N}\right) \to 0,$$

$$x\sum_{N}^{\infty}\frac{\sin^2(n/x)}{n^2}\log\frac{n}{N} = O\left(x\sum_{N}^{\infty}\frac{1}{n^2}\log\frac{n}{N}\right)$$

$$= O\left(x\int_{N}^{\infty}\frac{1}{t^2}\log\frac{t}{N}\,dt\right) = O\left(\frac{x}{N}\int_{1}^{\infty}\frac{\log u}{u^2}\,du\right) \to 0,$$

when $x/M \to \infty$ and $N/x \to \infty$, so that the conditions of Theorem 238 are satisfied. Thus, combining the theorem with Theorems 234, 70, and 68,† we obtain

THEOREM 240. *If $s_n \to s\,(\mathrm{R}_2)$, and s_n is real and slowly decreasing, then $s_n \to s$.*

12.15. Borel summability. We proved in § 9.13 that $s_n \to s$ (B) *and* $a_n = O(n^{-\frac{1}{2}})$ *imply* $s_n \to s$. Our object now is to prove

THEOREM 241. *If $s_n \to s$ (B) and*

(12.15.1) $$\varliminf (s_n - s_m) \geqslant 0$$

when

(12.15.2) $$m \to \infty, \qquad n > m, \qquad \frac{n-m}{\sqrt{m}} \to 0,$$

then $s_n \to s$.

This theorem, first proved by R. Schmidt and Vijayaraghavan, is the most general Tauberian theorem concerning (B) summability.

There are various methods. We may combine Theorem 238 with the ideas used by Hardy and Littlewood in their original proof of Theorem 156: this is the method followed by Vijayaraghavan. Alternatively, we may combine it with those used in §§ 9.10–13.‡ It is more natural here to combine Theorem 238 with a theorem of Wiener's type, and this is the course which we shall follow.

We observe first that it is unnecessary to distinguish between summability (B) and summability (B′). For $s_n \to s$ (B′) is equivalent to $s_{n-1} \to s$ (B), by Theorem 126, and the condition (12.15.1) is plainly unaltered by the change of m, n into $m-1$, $n-1$.

† See § 12.10(4).

‡ Certain parts of the argument of §§ 9.10–13, in which we used the full hypothesis $a_n = O(n^{-\frac{1}{2}})$, must then be modified, but the modifications required are not difficult.

We next verify that the (B) method of summation satisfies the conditions of Theorem 238, with $\phi(u) = 2\sqrt{u}, \Phi(u) = \sqrt{u}-1$. We have to show that

(12.15.3) $$e^{-x}\sum_{0}^{M}\frac{x^n}{n!} \to 0,$$ (12.15.4) $$e^{-x}\sum_{N}^{\infty}\frac{x^n}{n!} \to 0,$$

(12.15.5) $$e^{-x}\sum_{N}^{\infty}(\sqrt{n}-\sqrt{N})\frac{x^n}{n!} \to 0,$$

when

(12.15.6) $$\sqrt{x}-\sqrt{M} \to \infty, \qquad \sqrt{N}-\sqrt{x} \to \infty.$$

If $0 \leqslant \sqrt{x}-\sqrt{M} = \mu$ then $M = \sqrt{M}(\sqrt{x}-\mu) \leqslant x-\mu\sqrt{x}$; and if $0 \leqslant \sqrt{N}-\sqrt{x} = \nu$ then $N \geqslant x+2\nu\sqrt{x} \geqslant x+\nu\sqrt{x}$. Hence (12.15.3) and (12.15.4) will follow from

$$\lim_{\mu\leqslant\sqrt{x},\, \mu\to\infty}\left(e^{-x}\sum_{n=0}^{x-\mu\sqrt{x}}\frac{x^n}{n!}\right) = 0, \qquad \lim_{x\to\infty,\, \nu\to\infty}\left(e^{-x}\sum_{x+\nu\sqrt{x}}^{\infty}\frac{x^n}{n!}\right) = 0;$$

and these are true by Theorem 137(4). As regards (12.15.5), we have

$$e^{-x}\sum_{n=N}^{\infty}(\sqrt{n}-\sqrt{N})\frac{x^n}{n!} < \frac{e^{-x}}{\sqrt{x}}\sum_{n=N}^{\infty}(n-N)\frac{x^n}{n!} \leqslant \frac{e^{-x}}{\sqrt{x}}\sum_{x+\nu\sqrt{x}}^{\infty}(n-x)\frac{x^n}{n!}$$

$$\leqslant \frac{e^{-x}}{\sqrt{x}}\sum_{\xi+\nu\sqrt{\xi}}^{\infty}(n-\xi)\frac{x^n}{n!} = \frac{e^{-x}}{\sqrt{x}}\sum_{\nu\sqrt{\xi}}^{\infty} m\frac{x^{\xi+m}}{(\xi+m)!},$$

where $\xi = [x]$. It follows from Theorem 137† that this is

$$O\left(\frac{1}{x}\sum_{\nu\sqrt{\xi}}^{\infty} me^{-m^2/2\xi}\right) = O\left(\frac{1}{x}\int_{\nu\sqrt{\xi}}^{\infty} te^{-t^2/2\xi}\,dt\right) = O\left(\int_{\frac{1}{2}\nu}^{\infty} ue^{-u^2}\,du\right) = o(1).$$

Thus conditions (i) and (ii) of Theorem 238 are satisfied. Hence (if the conditions of Theorem 241 are satisfied) s_n is bounded.

In order to use Wiener's theorems, we must express summability (B) by an integral relation of Wiener's type. We may suppose, after what precedes, that s_n is bounded, and we may take $s = 0$, so that $s_n \to 0$ (B). Then we proved in § 9.10 (Theorem 151) that

$$\frac{1}{\sqrt{x}}\int_{0}^{\infty} e^{-(t-x)^2/2x}s(t)\,dt \to 0$$

when $x \to \infty$,‡ i.e. that

(12.15.7) $$\int_{0}^{\infty}\exp\left\{-\frac{(u^2-y^2)^2}{2y^2}\right\}\frac{u}{y}s(u^2)\,du \to 0$$

† Using (9.1.8) for the range $(\nu\sqrt{\xi}, \xi^{\zeta})$ and (9.1.6) for (ξ^{ζ}, ∞). See the remark in § 9.10 about the triviality of the 'tails' of such sums.

‡ Actually we assumed only that $s_n = o(\sqrt{n})$.

when $y \to \infty$. We now replace this by the simpler formula

(12.15.8) $$\int e^{-2(u-y)^2} s(u^2)\, du \to 0.$$

It is plain that (12.15.8) will follow from (12.15.7), when $s(t)$ is bounded, if

(12.15.9) $$\int \left| e^{-2(u-y)^2} - \frac{u}{y} \exp\left\{-\frac{(u^2-y^2)^2}{2y^2}\right\} \right| du \to 0,$$

i.e. if

(12.15.10) $$J = \int_{-y}^{\infty} |\phi(w,y)|\, dw \to 0,$$

where $$\phi(w,y) = e^{-2w^2} - \frac{y+w}{y} \exp\left\{-2w^2 \left(\frac{2y+w}{2y}\right)^2\right\}.$$

We divide J into the two parts J_1 and J_2 in which $|w| \geqslant y^\alpha$ and $|w| \leqslant y^\alpha$, where $0 < 3\alpha < 1$. In J_1, for large y,

$$0 < \frac{y+w}{y} < |w|, \qquad \frac{2y+w}{2y} > \tfrac{1}{2},$$

and $\phi = O(|w| e^{-\frac{1}{2}w^2})$, so that J_1 is trivial. In J_2 we have $w = O(y^\alpha)$,

$$\frac{y+w}{y} = 1 + O(y^{\alpha-1}), \qquad \left(\frac{2y+w}{2y}\right)^2 = 1 + O(y^{\alpha-1}),$$

$$\exp\left\{-2w^2 \left(\frac{2y+w}{2y}\right)^2\right\} = \exp\{-2w^2 + O(y^{3\alpha-1})\} = e^{-2w^2}\{1 + O(y^{3\alpha-1})\},$$

$$J_2 = O(y^{3\alpha-1}) \int e^{-2w^2}\, dw \to 0.$$

This proves (12.15.9) and therefore (12.15.8).

We now take $g(t) = e^{-2t^2}$ and $f(t) = s(t^2)$ for $t \geqslant 0$, 0 for $t < 0$. Then g is W (§ 12.9(1)). If $t > u$, $u \to \infty$, $t-u \to 0$, $\tau = t^2$, $v = u^2$, then

$$\frac{\tau - v}{\sqrt{v}} = \frac{t^2 - u^2}{u} < 2(t-u) \to 0,$$

and so $$\underline{\lim}\{f(t) - f(u)\} = \underline{\lim}\{s(\tau) - s(v)\} \geqslant 0.$$

Hence $f(t)$ is slowly decreasing (§ 12.2); and, since we have proved it bounded, it follows from Theorem 221 that $f(t) \to 0$, i.e. that $s_n \to 0$.

12.16. Summability (R, 2). We end this chapter by proving the theorem for (R, 2) summability which corresponds to Theorems 106 and 240, viz.

THEOREM 242. *If $s_n \to s$ (R, 2), i.e. if*

$$\chi(\theta) = \sum a_n \left(\frac{\sin n\theta}{n\theta}\right)^2 \to s$$

when $\theta \to 0$, and s_n is slowly decreasing, then $\sum a_n = s$.

This theorem, and its analogue for $f(t)$, present fresh difficulties, since the $g(t)$ now relevant, as we saw in §§ 12.10(5) and § 12.12, is not positive. If we knew that s_n is bounded, we could prove the theorem on the lines sketched in § 12.10, first deducing summability (C, k) from our general theorems, and then passing to convergence by means of Theorems 70 and 68. But we cannot supplement this argument here by an appeal to our later theorems, and it seems simplest to use different methods.

We write

$$a_n^+ = a_n \ (a_n \geqslant 0), \quad 0 \ (a_n < 0); \qquad a_n^- = a_n \ (a_n \leqslant 0), \quad 0 \ (a_n > 0);$$

so that $a_n = a_n^+ + a_n^-$, $|a_n| = a_n^+ - a_n^-$. We are given that $\underline{\lim}(s_n - s_m) \geqslant 0$ when $n > m$, $m \to \infty$, $(n-m)/m \to 0$; and it follows, taking $m = n-1$, that $a_n^- \to 0$, and that $\sum n^{-2} a_n^-$ is (absolutely) convergent.

Next, $\sum n^{-2} a_n \sin^2 n\theta$ is (by hypothesis) convergent for small θ. It follows that $\sum n^{-2} a_n^+ \sin^2 n\theta$ is convergent for small θ, and therefore, by Egoroff's theorem, uniformly convergent in a set E of positive measure mE. Hence

$$\sum \frac{a_n^+}{n^2} \int\limits_E \sin^2 n\theta \, d\theta < \infty.$$

But

$$\int\limits_E \sin^2 n\theta \, d\theta = \tfrac{1}{2} \int\limits_E (1 - \cos 2n\theta) \, d\theta \to \tfrac{1}{2} mE$$

when $n \to \infty$. Hence $\sum n^{-2} a_n^+$ is convergent, and $\sum n^{-2} a_n$ absolutely convergent.

We suppose, as we may, that $s = 0$. Then the series for $\theta^2 \chi(\theta)$ converges absolutely and uniformly for all positive θ, and $\chi(\theta) = o(1)$ when $\theta \to 0$. If $\delta > 0$ then

$$\int \frac{\delta \theta^2 \chi(\theta)}{\delta^2 + \theta^2} \, d\theta = \frac{1}{2} \sum \frac{a_n}{n^2} \int (1 - \cos 2n\theta) \frac{\delta \, d\theta}{\delta^2 + \theta^2} = \tfrac{1}{4}\pi \sum \frac{a_n}{n^2} (1 - e^{-2n\delta}),$$

the term-by-term integration being justified because

$$\sum \frac{|a_n|}{n^2} \int \frac{\delta \, d\theta}{\delta^2 + \theta^2} < \infty.$$

Differentiating twice with respect to δ, we find

$$\sum a_n e^{-2n\delta} = -\frac{1}{\pi} \int \theta^2 \chi(\theta) \frac{\partial^2}{\partial \delta^2} \left(\frac{\delta}{\delta^2 + \theta^2} \right) d\theta = -\frac{2}{\pi} \int \theta^2 \chi(\theta) \frac{\delta(\delta^2 - 3\theta^2)}{(\delta^2 + \theta^2)^3} \, d\theta.$$

Since $\chi(\theta) = o(1)$, the integral here is

$$o\left\{ \int \frac{\delta \theta^2 |\delta^2 - 3\theta^2|}{(\delta^2 + \theta^2)^3} \, d\theta \right\} = o\left\{ \int \frac{t^2 |1 - 3t^2|}{(1 + t^2)^3} \, dt \right\} = o(1)$$

when $\delta \to 0$; and so $\sum a_n e^{-2n\delta} \to 0$, i.e. $s_n \to s$ (A). The conclusion now follows from Theorem 106.

We have proved that *if* $s_n \to s$ (R, 2), *and* s_n *is slowly decreasing*,† *then* $s_n \to s$ (A); and this is enough for our purpose. It is, however, an imperfect theorem because, as we stated in § 12.10(5) and shall prove in Appendix III, $s_n \to s$ (R, 2) implies $s_n \to s$ (A) without any restriction on s_n. If this were granted, then Theorem 242 would naturally be a direct corollary of Theorem 106.

NOTES ON CHAPTER XII

§ 12.1. Wiener's original investigations are contained in his book *The Fourier integral* and his paper 'Tauberian theorems', *Annals* (2), 33 (1932), 1–100: the latter includes an elaborate bibliography of earlier work. A good many generalizations and simplifications have been made since by other writers, particularly by Pitt, *PLMS* (2), 44 (1938), 243–88. An important intermediate paper is that of Bochner, *BS* (1933), 126–44: Bochner shows that Wiener's analysis may be simplified considerably if we are prepared to impose rather stronger conditions on the kernels g.

There is a very clear account of the theory in Widder, ch. 5: both his account and that in §§ 12.1–8 are based largely on Pitt's.

§ 12.3. The theorems stated without proof in this section will be found in Titchmarsh's *Fourier integrals*.

§ 12.4. For 'strong convergence' see Titchmarsh, *Theory of functions*, 386 et seq.; Littlewood, 45 et seq.

§ 12.9. All these kernels appear in Wiener's work except (8), which was introduced by Ingham, *JLMS*, 20 (1945), 171–80.

§ 12.10. The theorems referred to in this section have been proved by different writers with different degrees of generality, and we do not give detailed references here. Broadly, the results of (1)–(3) were known before Wiener, his contribution being to include them in his general theory, while in (4) and (5) the results also are mostly his.

The papers of Szász and of Hardy and Rogosinski referred to will be found in *AM*, 61 (1933), 185–201 and *JLMS*, 18 (1943), 50–7.

§ 12.11. We have given Ingham's proof of the prime number theorem, l.c. under § 12.9. Wiener's proof is based on the kernel (7) (of § 12.9). Cf. Widder, 224–33. Widder includes proofs that $\zeta(1+i\tau) \neq 0$ and of all arithmetical theorems needed. Wiener and Widder aim directly at (12.11.3). If we choose, as in the text, to work with $\mu(n)$, then it is convenient to take

(*a*) $\sum n^{-1}\mu(n)$ *converges to* 0

as our immediate goal: (12.11.4) is a corollary, by Theorem 26 (§ 4.7). If

$$f(t) = \sum_{n \leqslant t} \frac{\mu(n)}{n},$$

† Actually we have *used* much less, in fact only $\underline{\lim}\, a_n \geqslant 0$.

then $f(t)$, being the sum-function of a series whose general term is $O(n^{-1})$, is slowly oscillating. Also, if $t \geqslant 1$,

$$\sum_{n\leqslant t} \mu(n)\left[\frac{t}{n}\right] = \sum_{n\leqslant t}\mu(n)\sum_{m\leqslant t/n} 1 = \sum_{mn\leqslant t}\mu(n) = \sum_{q\leqslant t}\sum_{n|q}\mu(n) = 1.$$

Hence
$$t\sum_{n\leqslant t}\frac{\mu(n)}{n} = 1+\sum_{n\leqslant t}\mu(n)\left(\frac{t}{n}-\left[\frac{t}{n}\right]\right) = O(t)$$

for $t \geqslant 1$, and so $f(t)$ (which is zero for $t < 1$) is bounded. Thus $f(t)$ satisfies the conditions of Theorem 233, and it is sufficient to prove (12.8.2) with this f and with g as in § 12.9(7).

If
$$F(y) = \sum \mu(n)\frac{e^{-ny}}{1-e^{-ny}},$$

then, on the one hand,

$$F(y) = \sum_n \mu(n)\sum_m e^{-mny} = \sum_q e^{-qy}\sum_{n|q}\mu(n) = e^{-y} = O(1) = o\left(\frac{1}{y}\right),$$

when $y \to 0$, and on the other,

$$F(y) = \frac{1}{y}\sum\frac{\mu(n)}{n}\,\frac{nye^{-ny}}{1-e^{-ny}}$$

$$= \frac{1}{y}\sum f(n)\Delta\frac{nye^{-ny}}{1-e^{-ny}} = \frac{1}{y}\sum f(n)\int_n^{n+1}\frac{d}{dt}\left(\frac{yte^{-yt}}{1-e^{-yt}}\right)dt = \int_0^\infty f(t)g(yt)\,dt.$$

Thus the last integral is $o(y^{-1})$, and this proves (12.8.2).

There is a further and quite different proof of the prime number theorem, based on Wiener's ideas as developed by Ikehara. For this see Widder, 233 et seq. The proof was reduced to its simplest form by Landau, *BS* (1932), 514–21. See also Bochner, *MZ*, 37 (1933), 1–9; Heilbronn and Landau, ibid. 10–16, 17, and 18–21; Karamata, *MZ*, 38 (1934), 701–8.

§ 12.12. Theorem 236 is apparently new. There is a theorem of Pitt [*DMJ*, 4 (1938), 437–40], proved in Widder, 215–21, which we might use instead and which imposes less restriction on g, but the proof is more difficult, and Theorem 236 is sufficient for our purposes here.

§§ 12.13–14. Vijayaraghavan, *JLMS*, 1 (1926), 113–20, and *PLMS* (2), 27 (1928), 316–26, proved the two cases of Theorem 238 required for A and B summability. The arguments which he used in these cases contain all the essential features of the proof of the general theorem. See also Karamata, *MZ*, 34 (1932), 737–40, and 37 (1933), 582–8.

§ 12.15. Theorem 241 was first proved in this form by R. Schmidt, *Schriften d. Königsberger gelehrten Gesellschaft*, 1 (1925), 205–56; and other proofs have been given by Vijayaraghavan, l.c. *supra*, and Wiener (l.c. under § 12.1). The proof here is essentially a simplification of Wiener's.

§ 12.16. For Egoroff's theorem see Titchmarsh, *Theory of functions*, 339, or Littlewood, 30–1.

XIII

THE EULER-MACLAURIN SUM FORMULA

13.1. Introduction. The 'Euler-Maclaurin sum formula'

$$(13.1.1)\quad \sum_{m=1}^{n} f(m) \sim \int_a^n f(x)\,dx + C + \tfrac{1}{2}f(n) + \sum_{r=1}^{\infty} (-1)^{r-1}\frac{B_r}{(2r)!} f^{(2r-1)}(n)$$

expresses the finite sum on the left in terms of the integral and the derivatives of $f(x)$. The exact theory of the formula belongs more to that of asymptotic than of summable series, but it is so important in many branches of analysis that we must discuss it seriously here.

We begin by considering one or two particularly simple cases. It is plain that the formula will be most useful when $f(x)$ behaves regularly for large x and the order of the kth derivative $f^{(k)}(x)$, considered as a function of x, decreases as k increases.

We suppose first that $0 < a \leqslant 1$, that $f'(x)$ is continuous for $x \geqslant a$, that $f > 0$, $f' < 0$, and that $f \to 0$ when $x \to \infty$. Then

$$\int_1^x |f'(t)|\,dt = -\int_1^x f'(t)\,dt = f(1) - f(x) \to f(1)$$

when $x \to \infty$, so that

$$\int_1^\infty |f'(t)|\,dt = f(1) < \infty.$$

If $y = x - [x]$, so that $y = x - m + 1$ for $m - 1 \leqslant x < m$, then $0 \leqslant y < 1$ and

$$J = \int_1^\infty y f'(x)\,dx$$

is absolutely convergent. Also

$$j_m = f(m) - \int_{m-1}^{m} f(x)\,dx = \int_{m-1}^{m} \{f(m) - f(x)\}\,dx$$

$$= \int_{m-1}^{m} \{f(m) - f(x)\}\frac{dy}{dx}\,dx = \int_{m-1}^{m} y f'(x)\,dx,$$

$$\sum_{m=2}^{n} f(m) - \int_1^n f(x)\,dx = \sum_{m=2}^{n} j_m = \int_1^n (x - [x]) f'(x)\,dx.$$

It follows that, if

$$(13.1.2)\qquad F(x) = \int_a^x f(t)\,dt,$$

then

$$\sum_{m=1}^{n} f(m)-F(n) \to f(1)-F(1)+J \tag{13.1.3}$$

when $n \to \infty$. If our conditions are satisfied for every $a > 0$, and $f(x)$ is integrable down to 0,† we may take $a = 0$ in (13.1.2) and (13.1.3).

Secondly, suppose that $f''(x)$ is continuous for $x \geqslant a$; that $f' > 0$, $f'' < 0$; and that $f' \to 0$ when $x \to \infty$. Then

$$\int_1^x |f''(t)|\,dt = -\int_1^x f''(t)\,dt = f'(1)-f'(x) \to f'(1)$$

and $|f''|$ is integrable up to ∞. Thus

$$J' = \tfrac{1}{2}\int_1^{\infty} (y^2-y)f''(x)\,dx$$

is absolutely convergent. If

$$j'_m = \tfrac{1}{2}\{f(m-1)+f(m)\}-\int_{m-1}^{m} f(x)\,dx$$

$$= f(m)-\int_{m-1}^{m} \{f(x)+\tfrac{1}{2}f'(x)\}\,dx = j_m-\tfrac{1}{2}\int_{m-1}^{m} f'(x)\,dx,$$

then

$$j''_m = \int_{m-1}^{m} (y-\tfrac{1}{2})f'(x)\,dx$$

$$= \frac{1}{2}\int_{m-1}^{m} f'(x)\frac{d(y^2-y)}{dx}\,dx = -\frac{1}{2}\int_{m-1}^{m} (y^2-y)f''(x)\,dx,$$

since $y^2-y \to 0$ when $x \to m-1+0$ or $x \to m-0$. Hence, summing from $m = 2$ to $m = n$,

$$\tfrac{1}{2}f(1)+f(2)+\ldots+f(n-1)+\tfrac{1}{2}f(n)-\int_1^n f(x)\,dx = -\tfrac{1}{2}\int_1^n (y^2-y)f''(x)\,dx.$$

It follows that

$$\sum_{m=1}^{n} f(m)-F(n)-\tfrac{1}{2}f(n) \to \tfrac{1}{2}f(1)-F(1)-J'. \tag{13.1.4}$$

We may regard (13.1.3) and (13.1.4) as the two simplest cases of (13.1.1), with

$$C = f(1)-F(1)+J, \qquad C = \tfrac{1}{2}f(1)-F(1)-J'$$

respectively. If, for example, $f(x) = \log x$ and $a = 1$, then our second set of conditions is satisfied, and we find that

$$\log n!-(n+\tfrac{1}{2})\log n+n \to A = 1+\frac{1}{2}\int_1^{\infty} \frac{y^2-y}{x^2}\,dx,$$

† In which case $xf \to 0$ when $x \to 0$ and xf' is also integrable down to 0.

or that $n! \sim e^A n^{n+\frac{1}{2}} e^{-n}$. This is (apart from the calculation of A, which is actually $\frac{1}{2}\log 2\pi$), the simplest form of Stirling's theorem, and it is natural to expect that a fuller investigation of the formula (13.1.1) will lead to a complete asymptotic expansion of $\log n!$.

13.2. The Bernoullian numbers and functions. We shall investigate the formula in this chapter by two different methods, by real analysis in §§ 13.5–7, and by the use of Cauchy's theorem in §§ 13.14–16: the second method will naturally demand much more stringent conditions on $f(x)$. Our first method depends upon the properties of the Bernoullian functions $B_n(x)$.

We define the Bernoullian numbers B_n, and the functions $B_n(x)$ and $\phi_n(x)$, by

(13.2.1)
$$\frac{t}{e^t-1} = 1-\tfrac{1}{2}t+B_1\frac{t^2}{2!}-B_2\frac{t^4}{4!}+\dots$$
$$= 1-\tfrac{1}{2}t+\sum(-1)^{n-1}B_n\frac{t^{2n}}{(2n)!},$$

(13.2.2)
$$t\frac{e^{xt}}{e^t-1} = 1+\sum B_n(x)\frac{t^n}{n!},$$

(13.2.3)
$$t\frac{e^{xt}-1}{e^t-1} = \sum \phi_n(x)\frac{t^n}{n!}.$$

Here, and to the end of the chapter, sums without limits run from 1 to ∞. The left-hand side of (13.2.3) is 0 for $x = 0$ and t for $x = 1$, so that

(13.2.4) $\phi_n(0) = 0;$ (13.2.5) $\phi_1(1) = 1,\quad \phi_n(1) = 0\ (n > 1).$

The series are convergent for $|t| < 2\pi$. The first B_n are

$$B_1 = \tfrac{1}{6},\quad B_2 = \tfrac{1}{30},\quad B_3 = \tfrac{1}{42},\quad B_4 = \tfrac{1}{30},\quad B_5 = \tfrac{5}{66},\quad \dots.$$

It is plain that

$$B_1(x) = \phi_1(x)-\tfrac{1}{2} = x-\tfrac{1}{2},\qquad B_2(x) = \phi_2(x)+\tfrac{1}{6},$$
$$B_3(x) = \phi_3(x),\qquad B_4(x) = \phi_4(x)-B_2,\quad \dots,$$

and generally

(13.2.6) $B_{2r}(x) = \phi_{2r}(x)+(-1)^{r-1}B_r,\qquad B_{2r+1}(x) = \phi_{2r+1}(x)\ (r > 0).$

In particular

(13.2.7)
$$B_{2r}(0) = B_{2r}(1) = (-1)^{r-1}B_r,\qquad B_{2r+1}(0) = B_{2r+1}(1) = 0\quad (r > 0).$$

It is familiar that

(13.2.8)
$$B_n = \frac{(2n)!}{2^{2n-1}\pi^{2n}}\sum\frac{1}{m^{2n}}$$

(so that B_n increases rapidly for large n).

The first $\phi_n(x)$ are x, x^2-x, $x^3-\frac{3}{2}x^2+\frac{1}{2}x,\dots$. Writing (13.2.3) as

$$\sum \phi_n(x)\frac{t^n}{n!} = \left(1-\tfrac{1}{2}t+B_1\frac{t^2}{2!}-\dots\right)\sum \frac{x^n t^n}{n!},$$

and equating coefficients, we find that

(13.2.9) $$\phi_n(x) = x^n-\tfrac{1}{2}nx^{n-1}+\binom{n}{2}B_1 x^{n-2}-\binom{n}{4}B_2 x^{n-4}+\dots,$$

the last term being one in x or x^2.

The identity $$t\frac{e^{(x+1)t}-1}{e^t-1}-t\frac{e^{xt}-1}{e^t-1} = te^{xt}$$

shows that

(13.2.10) $$\phi_n(x+1)-\phi_n(x) = nx^{n-1},$$

and from this and (13.2.4) it follows that

(13.2.11) $$1^{n-1}+2^{n-1}+\dots+N^{n-1} = n^{-1}\phi_n(N+1) \quad (n > 1).$$

Differentiating (13.2.2) with respect to x, we find

$$\sum B_n'(x)\frac{t^n}{n!} = \frac{t^2e^{xt}}{e^t-1} = t+\sum B_n(x)\frac{t^{n+1}}{n!},$$

and hence, equating coefficients,

(13.2.12) $$B_1'(x) = 1, \quad B_n'(x) = nB_{n-1}(x) \quad (n > 1).$$

The corresponding equations for the $\phi_n(x)$ are

$$\phi_1'(x) = 1, \qquad \phi_2'(x) = 2\{\phi_1(x)-\tfrac{1}{2}\}, \qquad \phi_3'(x) = 3\{\phi_2(x)+B_1\}, \quad \dots$$

and generally

(13.2.13)

$$\phi_{2m}'(x) = 2m\phi_{2m-1}(x), \qquad \phi_{2m+1}'(x) = (2m+1)\{\phi_{2m}(x)+(-1)^{m-1}B_m\},$$

the first for $m = 2, 3,\dots$, the second for $m = 1, 2,\dots$.

13.3. The associated periodic functions. We now define $\mathrm{B}_n(x)$ and $\psi_n(x)$ as the functions equal to $B_n(x)$ and $\phi_n(x)$ for $0 \leqslant x < 1$ and with period 1. It follows from (13.2.4)–(13.2.6) that $\mathrm{B}_n(x)$ and $\psi_n(x)$ are continuous for all x if $n > 1$, while $\mathrm{B}_1(x)$ and $\psi_1(x)$ have a jump -1 for every integral x.

We know that

$$\frac{1}{\pi}\sum \frac{\sin 2m\pi x}{m} = \tfrac{1}{2}-x = -\{\phi_1(x)-\tfrac{1}{2}\}$$

for $0 < x < 1$, and that the series is boundedly convergent. If we integrate term by term, we obtain

$$\frac{1}{2\pi^2}\sum \frac{1-\cos 2m\pi x}{m^2} = \tfrac{1}{2}x-\tfrac{1}{2}x^2 = -\tfrac{1}{2}\phi_2(x),$$

and so $$\frac{1}{2\pi^2}\sum \frac{\cos 2m\pi x}{m^2} = \tfrac{1}{2}(x^2-x+\tfrac{1}{6}) = \tfrac{1}{2}\{\phi_2(x)+B_1\}.$$

Hence

(13.3.1) $$\frac{1}{2\pi^2}\sum \frac{\cos 2m\pi x}{m^2} = \tfrac{1}{2}\{\psi_2(x)+B_1\} = \tfrac{1}{2}\mathrm{B}_2(x)$$

for all x. Generally

(13.3.2)
$$\frac{1}{2^{2k-1}\pi^{2k}}\sum \frac{\cos 2m\pi x}{m^{2k}} = \frac{(-1)^{k-1}}{(2k)!}\{\psi_{2k}(x)+(-1)^{k-1}B_k\} = \frac{(-1)^{k-1}}{(2k)!}\mathrm{B}_{2k}(x)$$

and

(13.3.3)
$$\frac{1}{2^{2k}\pi^{2k+1}}\sum \frac{\sin 2m\pi x}{m^{2k+1}} = \frac{(-1)^{k-1}}{(2k+1)!}\psi_{2k+1}(x) = \frac{(-1)^{k-1}}{(2k+1)!}\mathrm{B}_{2k+1}(x),$$

for $k = 1, 2,\ldots$ and all x. For (13.2.7) shows that they are true for $x = 0$; (13.3.2) is the formal derivative of (13.3.3); and (13.3.3), with $k-1$ for k and the sign changed, is the formal derivative of (13.3.2). Finally, (13.3.3) becomes true for $k = 0$ and non-integral x if we substitute $\psi_1^*(x) = \psi_1(x)-\frac{1}{2}$ for $\psi_1(x)$.

13.4. The signs of the functions $\phi_n(x)$. It follows from (13.2.4) and (13.2.5) that all the ϕ_n after ϕ_1 vanish for $x = 0$ and $x = 1$. We now prove

THEOREM 243. *The functions $\phi_2, \phi_4, \phi_6,\ldots$ have fixed signs in $(0, 1)$, that of ϕ_{2k} being $(-1)^k$, while $\phi_3, \phi_5,\ldots$ vanish also for $x = \frac{1}{2}$, and ϕ_{2k+1} has the signs $(-1)^{k-1}$ and $(-1)^k$ in $(0, \frac{1}{2})$ and in $(\frac{1}{2}, 1)$ respectively.*

First,

$$\sum \phi_n(\tfrac{1}{2})\frac{t^n}{n!}-\sum \phi_n(\tfrac{1}{2})\frac{(-t)^n}{n!} = \frac{t}{e^{\frac{1}{2}t}+1}+\frac{t}{e^{-\frac{1}{2}t}+1} = t,$$

so that $\phi_3(\frac{1}{2}) = \phi_5(\frac{1}{2}) = \ldots = 0$. Thus $\phi_{2k+1}(x)$ has the three zeros $0, \frac{1}{2}, 1$.

Next, our assertion is true of

$$\phi_2(x) = x(x-1), \qquad \phi_3(x) = x(x-\tfrac{1}{2})(x-1).$$

We assume that it is true up to ϕ_{2m-1}, and prove that it is true of ϕ_{2m} and ϕ_{2m+1}. Since $\phi'_{2m} = 2m\phi_{2m-1}$ vanishes at 0, $\frac{1}{2}$, and 1 only, ϕ_{2m} is

of fixed sign in $(0, 1)$; and (13.2.9), for small x, shows that the sign is $(-1)^m$. Also ϕ_{2m} is monotone in $(0, \frac{1}{2})$ and in $(\frac{1}{2}, 1)$, and so

$$\phi'_{2m+1}(x) = (2m+1)\{\phi_{2m}(x)+(-1)^{m-1}B_m\}$$

can vanish at most once in each of these intervals. Thus ϕ_{2m+1}, which vanishes at 0, $\frac{1}{2}$, and 1, is of fixed sign in $(0, \frac{1}{2})$ and in $(\frac{1}{2}, 1)$, its sign in the first of these intervals being $(-1)^{m-1}$, by (13.2.9), and in the second $(-1)^m$, since $\phi'_{2m+1}(\frac{1}{2}) \neq 0$.

We shall also use the properties

(13.4.1) $$B_{2m-1}(x) = -B_{2m-1}(1-x), \quad B_{2m}(x) = B_{2m}(1-x)$$
$$(0 < x < 1;\ m = 1, 2, \ldots).$$

These equations follow from the trigonometrical developments of § 13.3, or from the fact that

$$t\frac{e^{xt}+e^{(1-x)t}}{e^t-1} = t\frac{\cosh(x-\frac{1}{2})t}{\sinh\frac{1}{2}t}, \qquad t\frac{e^{xt}-e^{(1-x)t}}{e^t-1} = t\frac{\sinh(x-\frac{1}{2})t}{\sinh\frac{1}{2}t}$$

are even and odd functions of t respectively.

13.5. The Euler-Maclaurin sum formula. In what follows we assume the continuity of all derivatives of $f(x)$ which occur for $x > 0$: $f(x)$ will usually have a singularity at $x = 0$. We define $F(x)$ as in (13.1.2): a will usually be taken to be 0 when $f(x)$ is integrable down to 0.

We suppose in the first instance that $0 \leqslant x \leqslant 1$, and that $f(x)$ and its derivatives are continuous in this closed interval; and we consider the integral

(13.5.1) $$\rho_r = \rho_r(x) = -\frac{1}{r!}\int_0^1 \mathrm{B}_r(x-t)f^{(r)}(t)\,dt,$$

where $r \geqslant 1$. We must distinguish the cases $r > 1$ and $r = 1$.

If $r > 1$ then $\mathrm{B}_r(u)$ is continuous, and

$$\frac{d\mathrm{B}_r(x-t)}{dt} = -r\mathrm{B}_{r-1}(x-t),$$

by (13.2.12). Also $\mathrm{B}_r(x-1) = \mathrm{B}_r(x) = B_r(x)$. Hence

(13.5.2)

$$\rho_r = -\frac{\mathrm{B}_r(x)}{r!}\{f^{(r-1)}(1)-f^{(r-1)}(0)\}-\frac{1}{(r-1)!}\int_0^1 \mathrm{B}_{r-1}(x-t)f^{(r-1)}(t)\,dt$$
$$= -\frac{B_r(x)}{r!}\{f^{(r-1)}(1)-f^{(r-1)}(0)\}+\rho_{r-1} \quad (r > 1).$$

If $r = 1$, we suppose first that $0 < x < 1$. In this case $\mathrm{B}_r(u)$ has jumps -1 for integral u, and differential coefficient 1 elsewhere, so that

$$\mathrm{B}_1(-0)-\mathrm{B}_1(+0) = 1, \qquad \frac{d\mathrm{B}_1(x-t)}{dt} = -1.$$

Hence

$$f(x)-\int_0^1 f(t)\,dt = \{\mathrm{B}_1(-0)-\mathrm{B}_1(+0)\}f(x)+\int_0^1 \frac{d\mathrm{B}_1(x-t)}{dt}f(t)\,dt.$$

Also

$$\int_0^x \frac{d\mathrm{B}_1(x-t)}{dt}f(t)\,dt = \mathrm{B}_1(+0)f(x)-\mathrm{B}_1(x)f(0)-\int_0^x \mathrm{B}_1(x-t)f'(t)\,dt,$$

$$\int_x^1 \frac{d\mathrm{B}_1(x-t)}{dt}f(t)\,dt = \mathrm{B}_1(x-1)f(1)-\mathrm{B}_1(-0)f(x)-\int_x^1 \mathrm{B}_1(x-t)f'(t)\,dt,$$

and $\mathrm{B}_1(x-1) = \mathrm{B}_1(x) = B_1(x)$. Hence, combining the last three equations, we obtain

$$(13.5.3)\quad f(x)-\int_0^1 f(t)\,dt = B_1(x)\{f(1)-f(0)\}-\int_0^1 \mathrm{B}_1(x-t)f'(t)\,dt$$

$$= B_1(x)\{f(1)-f(0)\}+\rho_1.$$

We have proved this for $0 < x < 1$, and it holds, by continuity, for $0 \leqslant x \leqslant 1$.

Supposing now that $l > 1$, and combining (13.5.3) with (13.5.2) for $r = 2, 3,\ldots, l$, we obtain

$$(13.5.4)\qquad f(x) = \int_0^1 f(t)\,dt+\sum_{r=1}^{l}\frac{B_r(x)}{r!}\{f^{(r-1)}(1)-f^{(r-1)}(0)\}+\rho_l$$

for $0 \leqslant x \leqslant 1$. The equation reduces to (13.5.3) for $l = 1$, so that it is true for $l \geqslant 1$.

We now replace $f(x)$ by $f(x+m-1)$, where m is a positive integer, and obtain

$$f(x+m-1) = \int_{m-1}^{m} f(t)\,dt+\sum_{r=1}^{l}\frac{B_r(x)}{r!}\{f^{(r-1)}(m)-f^{(r-1)}(m-1)\}+\rho_{l,m},$$

where $$\rho_{l,m} = -\frac{1}{l!}\int_{m-1}^{m} \mathrm{B}_l(x-t)f^{(l)}(t)\,dt.$$

If, in particular, we take $x = 1$, $l = 2k+1$, observe that

$$B_1(1) = \tfrac{1}{2}, \qquad B_{2s-1}(1) = 0 \quad (s > 1), \qquad B_{2s}(1) = (-1)^{s-1}B_s,$$
$$\mathrm{B}_{2k+1}(1-t) = -\mathrm{B}_{2k+1}(t) = -\psi_{2k+1}(t),$$

by (13.2.7) and (13.4.1), and then replace s by r, we obtain

$$f(m) = \int_{m-1}^{m} f(t)\,dt + \tfrac{1}{2}\{f(m)-f(m-1)\} + S_k(m) - S_k(m-1) + \sigma_{k,m}$$

or

(13.5.5) $$\tfrac{1}{2}\{f(m-1)+f(m)\} = \int_{m-1}^{m} f(t)\,dt + S_k(m) - S_k(m-1) + \sigma_{k,m},$$

where

(13.5.6) $$S_k(m) = \sum_{r=1}^{k} \frac{(-1)^{r-1}B_r}{2r!} f^{(2r-1)}(m),$$

(13.5.7) $$\sigma_{k,m} = \frac{1}{(2k+1)!} \int_{m-1}^{m} \psi_{2k+1}(t) f^{(2k+1)}(t)\,dt$$
$$= -\frac{1}{(2k+2)!} \int_{m-1}^{m} \psi_{2k+2}(t) f^{(2k+2)}(t)\,dt,$$

by another partial integration, since $\psi_{2k+2}(0) = \psi_{2k+2}(1) = 0$. If $k = 0$ then the terms $S_k(m)$ and $S_k(m-1)$ disappear from (13.5.5).

Summing (13.5.5) for $m = 2, 3, \dots, n$, and adding $\tfrac{1}{2}f(1)+\tfrac{1}{2}f(n)$, we obtain

(13.5.8) $$\sum_{1}^{n} f(m) = F(n) + \tfrac{1}{2}f(n) + S_k(n) + P_k + U_{k,n},$$

(13.5.9) $$P_k = -F(1) + \tfrac{1}{2}f(1) - S_k(1),$$

(13.5.10) $$U_{k,n} = -\frac{1}{(2k+2)!} \int_{1}^{n} \psi_{2k+2}(t) f^{(2k+2)}(t)\,dt.$$

If we write

(13.5.11) $$\chi_{2k+2}(t) = \sum_{m=1}^{\infty} \frac{\cos 2m\pi t}{m^{2k+2}}$$
$$= (-1)^k \frac{2^{2k+1}\pi^{2k+2}}{(2k+2)!} \{\psi_{2k+2}(t) + (-1)^k B_{k+1}\},$$

then

$$U_{k,n} = \frac{(-1)^{k+1}}{2^{2k+1}\pi^{2k+2}} \int_{1}^{n} \chi_{2k+2}(t) f^{(2k+2)}(t)\,dt + \frac{(-1)^k B_{k+1}}{(2k+2)!} \int_{1}^{n} f^{(2k+2)}(t)\,dt$$
$$= \frac{(-1)^{k+1}}{2^{2k+1}\pi^{2k+2}} \int_{1}^{n} \chi_{2k+2}(t) f^{(2k+2)}(t)\,dt + \frac{(-1)^k B_{k+1}}{(2k+2)!} \{f^{(2k+1)}(n) - f^{(2k+1)}(1)\}.$$

Hence we may also write (13.5.8)–(13.5.10) as

(13.5.12) $$\sum_1^n f(m) = F(n)+\tfrac{1}{2}f(n)+S_{k+1}(n)+Q_k+V_{k,n},$$

(13.5.13) $$Q_k = -F(1)+\tfrac{1}{2}f(1)-S_{k+1}(1),$$

(13.5.14) $$V_{k,n} = \frac{(-1)^{k+1}}{2^{2k+1}\pi^{2k+2}}\int_1^n \chi_{2k+2}(t)f^{(2k+2)}(t)\,dt.$$

When $f(x)$ is a polynomial, $S_k(n)$ is also one, and $U_{k,n}$ vanishes for sufficiently large k. If, for example, $f(x) = x^l$, we may take k to be $\frac{1}{2}l$ or $\frac{1}{2}(l-1)$. In this case it is easily verified that (13.5.8) reduces to (13.2.11).

13.6. Limits as $n \to \infty$. So far there has been no question of convergence. We now introduce the hypotheses that

(13.6.1) $$\int^\infty |f^{(2k+2)}(x)|\,dx < \infty$$

and

(13.6.2) $$f^{(2k+1)}(x) \to 0$$

when $x \to \infty$; and write the integrals over $(1, n)$ as differences of integrals over $(1, \infty)$ and (n, ∞). We then find

(13.6.3) $$\sum_1^n f(m) = F(n)+\tfrac{1}{2}f(n)+S_k(n)+C_k+R_{k,n},$$

(13.6.4) $$C_k = -F(1)+\tfrac{1}{2}f(1)-S_k(1)-\frac{1}{(2k+2)!}\int_1^\infty \psi_{2k+2}(t)f^{(2k+2)}(t)\,dt,$$

(13.6.5) $$R_{k,n} = \frac{1}{(2k+2)!}\int_n^\infty \psi_{2k+2}(t)f^{(2k+2)}(t)\,dt.$$

There are alternative forms corresponding to (13.5.12)–(13.5.14).

It is plain, since $\psi_{2k+2}(t) = O(1)$, that $R_{k,n} \to 0$ when $n \to \infty$. Hence it follows from (13.6.3) that

(13.6.6) $$\sum_1^n f(m)-F(n)-\tfrac{1}{2}f(n)-S_k(n) \to C_k.$$

The most interesting case is that in which (13.6.1) and (13.6.2) are true for all k from a certain K. Then (13.6.6) is true for $k = K$ and $k = K+1$, and

$$S_{K+1}(n)-S_K(n) = \frac{(-1)^K B_{K+1}}{(2K+2)!}f^{(2K+1)}(n) \to 0.$$

It follows that $C_{K+1} = C_K$, so that C_k is independent of k for $k \geqslant K$. This is easily verified directly, since

$$-\frac{1}{(2k+2)!}\int\limits_1^\infty \psi_{2k+2}(t)f^{(2k+2)}(t)\,dt$$
$$= \frac{(-1)^{k-1}}{2k!}B_k f^{(2k-1)}(1) - \frac{1}{2k!}\int\limits_1^\infty \psi_{2k}(t)f^{(2k)}(t)\,dt,$$

by two partial integrations.

Thus C_k is, for $k \geqslant K$, independent of (n and) k, and $C = C_k$ is a number depending only on $f(x)$ and $F(x)$, i.e. on $f(x)$ and the lower limit a in (13.1.2). We call C the *Euler-Maclaurin constant* of f (and F). We shall also call C the $(\mathfrak{R}, a)$ sum of the series $\sum f(n)$, and write

(13.6.7) $$f(1)+f(2)+...+f(x)+... = C \quad (\mathfrak{R}, a).$$

We thus obtain another definition of the sum of a divergent series, but one of a quite different type from most of those which we have considered, and primarily adapted to series of positive terms such as $1+1+1+...$ or $\log 2+\log 3+\log 4+...$.

The $\mathfrak{R}$ stands for Ramanujan, whose work with divergent series was mainly based on this definition. The definition is implicit in much of Euler's work. The sum which it attributes to a series depends on the value chosen for a. We shall, however, find that there is usually one value of a which it is natural to choose in any special case.

We shall call (13.5.8), (13.6.3), or one or other of their variants, according to the context, 'the Euler-Maclaurin sum formula'.

13.7. The sign and magnitude of the remainder term. We now strengthen our hypotheses by supposing the derivatives of $f(x)$, from a certain point onwards, of constant sign. To fix our ideas, we suppose (13.6.1) and (13.6.2) true, and $f^{(2k+2)}(x) \leqslant 0$, for $k \geqslant K$ and $x \geqslant 1$.† If now $k > K$, then, after Theorem 243, $R_{k,n}$ has the sign $(-1)^k$ and $R_{k-1,n}$ the sign $(-1)^{k-1}$; so that $|R_{k,n}| \leqslant |R_{k-1,n} - R_{k,n}|$. But

$$R_{k,n} = -\frac{1}{(2k+2)!}\int f^{(2k+1)}\psi'_{2k+2}\,dt = -\frac{1}{(2k+1)!}\int f^{(2k+1)}\psi_{2k+1}\,dt$$
$$= \frac{1}{(2k+1)!}\int f^{(2k)}\psi'_{2k+1}\,dt = \frac{1}{2k!}\int \{\psi_{2k}+(-1)^{k-1}B_k\}f^{(2k)}\,dt$$

(all the integrations being from n to ∞), and so

$$R_{k-1,n} - R_{k,n} = \frac{(-1)^{k-1}B_k}{2k!}f^{(2k-1)}(n),$$

† In which case $f^{(2k+1)}(x)$ is non-negative for $k \geqslant K$, and $f^{(n)}(x) \to 0$ for $n \geqslant 2K+1$.

which is the last term of $S_k(n)$. Hence we obtain

THEOREM 244. *If $f(x)$ satisfies* (13.6.1) and (13.6.2), *and $f^{(2k+2)}(x)$ is of fixed sign, for $k \geqslant K$, then the error in the formula*

$$\sum_1^n f(m) = F(n)+\tfrac{1}{2}f(n)+S_k(n)+C$$

alternates in sign as k increases from $K+1$ onwards, and does not exceed the last term retained in the series; and the series (13.6.4) *for $C_k = C$ has the same properties.*

The series

(13.7.1) $$S(n) = \sum \frac{(-1)^{r-1}B_r}{2r!}f^{(2r-1)}(n)$$

and $S(1)$, obtained by making k infinite in $S_k(n)$ and $S_k(1)$, are usually divergent, owing to the rapid increase of B_r for large r. They may however, often be used effectively for purposes of numerical computation. If

(i) a_r is real,

(ii) $s = a_1+a_2+...+a_r+R_r$ for every r,

(iii) R_r alternates in sign,

then we may say that the series $\sum a_r$ *alternates round* s; we may suppose if we please that condition (iii) is satisfied only for $r \geqslant r_0$. It is plain that

$$|R_r| \leqslant |a_r|$$

(for $r > 1$ or $r > r_0$). The definition does not determine a unique s; if, for example, $R_{2r-1} < 0$ and $R_{2r} > 0$, and

$$\rho_1 = \min_r |R_{2r}|, \qquad \rho_2 = \min_r |R_{2r-1}|,$$

then $\sum a_r$ also alternates round any number of the interval $(s-\rho_1, s+\rho_2)$.†

The important case is that in which a_r, R_r, and s are functions of a parameter x, and

$$|R_r(x)| \to 0$$

as $x \to \infty$ (for $r > 1$ or $r > r_0$), as, for example, when

$$\sum a_r(x) = c_1+c_2x^{-1}+c_3x^{-2}+...$$

is a divergent asymptotic series for a function $g(x)$. If $\sum a_r(x)$ alternates round $g(x)$, in the sense just explained, then we shall say that $\sum a_r(x)$ is a *semi-convergent series for* $g(x)$. Thus our conditions are satisfied, with $x = n$, by the series (13.7.1), under the conditions of Theorem 244.

† Thus all the conditions are satisfied by the series $1-2+2-2+...$ with $s = 0$. Here R_r is alternately -1 and 1, and the series alternates round any number of $(-1, 1)$.

If r is given, $|R_r(x)| \to 0$ when $x \to \infty$. If x is given, $|R_r(x)|$ will usually tend to infinity when $r \to \infty$. It will, however, generally happen that, for a given x, $|R_r(x)|$ is conveniently small for a suitably chosen r, for example, for an r for which $|a_r(x)|$ takes its minimum m_r; and then the series may be used for the computation of $g(x)$. The computation will be the more accurate the larger x.

In these circumstances we may reasonably say that $\sum c_r = \sum a_r(1)$ is a semi-convergent series for $s = g(1)$. We cannot say that s is the 'sum' of the series, since $\sum c_r$ alternates round any number in an interval $(s-\rho_1, s+\rho_2)$; but s will often be the sum of the series in some other sense.† Further if $\sum a_r(x)$ is a semi-convergent series for $g(x)-h(x)$, we may say that

$$h(x)+\sum a_r(x)$$

is a semi-convergent series for $g(x)$.

For example, returning to the series (13.7.1), let us suppose that $f(x) = \log x$ and $a = 0$, so that $F(x) = x\log x - x$. Then our conditions are satisfied for $r \geqslant 1$, and we are led to the formulae

$$\log n! = \sum_1^n \log m = (n+\tfrac{1}{2})\log n - n + C + \frac{B_1}{1.2}\frac{1}{n} - \frac{B_2}{3.4}\frac{1}{n^3} + \frac{B_3}{5.6}\frac{1}{n^5} - \dots, \tag{13.7.2}$$

$$C = 1 - \frac{B_1}{1.2} + \frac{B_2}{3.4} - \frac{B_3}{5.6} + \frac{B_4}{7.8} - \dots. \tag{13.7.3}$$

The series are semi-convergent, and can be used to calculate $\log n!$ and C. We shall see later that $C = \frac{1}{2}\log 2\pi$.

We cannot calculate C with great accuracy from (13.7.3) because $n = 1$ is too small. The least term is that last written, which is $-\cdot 00059$, and we can calculate $C = \cdot 919...$, to 3 places, by stopping there. This value of C, used in (13.7.2), would then give a fairly accurate value for $\log n!$ for large n. On the other hand, we could calculate C, with much greater accuracy, by using (13.7.2) with a fairly large n and computing $\log n!$ independently.

In practice the C of a given f would be computed by writing

$$\sum f(n) = f(1)+\dots+f(N)+\sum f(n+N),$$

and applying our formulae to the last series, for which they will be more effective the larger N. A judicious choice of N should then make both parts of the calculation practicable with considerable accuracy.

The method may be applied to convergent series whose convergence is inconveniently slow. In this case we must take $a = \infty$, so that $F(n) \to 0$, and C is the sum of the series. Thus Euler, taking $f(x) = (x+9)^{-2}$, calculated

$$\frac{\pi^2}{6} = \frac{1}{1^2}+\frac{1}{2^2}+\dots+\frac{1}{9^2}+\sum\frac{1}{(n+9)^2}$$

to 18 places of decimals.

† See, for example, §§ 13.15–16.

13.8. Poisson's proof of the Euler-Maclaurin formula. Poisson, in his investigation of the formula, starts from the theory of Fourier series.

Suppose that $f(x)$ is indefinitely differentiable for $x > 0$, and that $0 < a < b < \infty$. Then, by the ordinary theory of Fourier series,

$$(13.8.1)\quad f(x) = \frac{1}{b-a}\int f(t)\,dt + \frac{2}{b-a}\sum\int f(t)\cos\frac{2\pi r(t-x)}{b-a}\,dt,$$

where the integrations are over (a,b), for $a < x < b$. For $x = a$ or $x = b$ the sum is $\frac{1}{2}\{f(a)+f(b)\}$. We take $n\omega = b-a$ and

$$x = a,\ a+\omega,\ a+2\omega,\ldots,\ a+(n-1)\omega,$$

substitute in (13.8.1), and add the results. We thus obtain

$$(13.8.2)\quad \tfrac{1}{2}f(a)+f(a+\omega)+\ldots+f\{a+(n-1)\omega\}+\tfrac{1}{2}f(b)$$
$$= \frac{n}{b-a}\int f(t)\,dt + \frac{2}{b-a}\sum_{r=1}^{\infty}\int f(t)\sum_{s=0}^{n-1}\cos\frac{2\pi r(t-a-s\omega)}{b-a}\,dt.$$

The sum under the integral sign is

$$\Re\left\{\exp\frac{2\pi ir(t-a)}{b-a}\sum_{s=0}^{n-1}\exp\left(-\frac{2\pi irs}{n}\right)\right\};$$

and the sum here is 0 unless $r = ln$, where l is a positive integer, and then it is n. Hence (13.8.2) is

(13.8.3)
$$\tfrac{1}{2}f(a)+f(a+\omega)+\ldots+\tfrac{1}{2}f(b) = \frac{1}{\omega}\int f(t)\,dt + \frac{2}{\omega}\sum_{l=1}^{\infty}\int f(t)\cos\frac{2\pi l(t-a)}{\omega}\,dt.$$

Now
$$\int f(t)\cos\frac{2\pi l(t-a)}{\omega}\,dt = -\frac{\omega}{2l\pi}\int f'(t)\sin\frac{2\pi l(t-a)}{\omega}\,dt$$
$$= \sum_{r=1}^{k}(-1)^{r-1}\left(\frac{\omega}{2l\pi}\right)^{2r}\{f^{(2r-1)}(b)-f^{(2r-1)}(a)\}+$$
$$+(-1)^{k}\left(\frac{\omega}{2l\pi}\right)^{2k}\int f^{(2k)}(t)\cos\frac{2\pi l(t-a)}{\omega}\,dt,$$

by repeated partial integration. Substituting in (13.8.3), and using (13.2.8), we obtain

$$(13.8.4)\quad \tfrac{1}{2}f(a)+f(a+\omega)+\ldots+\tfrac{1}{2}f(b)$$
$$= \frac{1}{\omega}\int f(t)\,dt + \sum_{r=1}^{k}(-1)^{r-1}\omega^{2r-1}\frac{B_r}{2r!}\{f^{(2r-1)}(b)-f^{(2r-1)}(a)\}+W_k,$$

where
$$W_k = (-1)^k\frac{\omega^{2k-1}}{2^{2k-1}\pi^{2k}}\int f^{(2k)}(t)\sum\frac{1}{l^{2k}}\cos\frac{2\pi l(t-a)}{\omega}\,dt.$$

If, in particular, we take $a = 1$, $\omega = 1$, $b = n+1$, then (13.8.4) becomes

$$f(1)+f(2)+...+f(n+1) = F(n+1)-F(1)+\tfrac{1}{2}f(n+1)+\tfrac{1}{2}f(1)+$$
$$+\sum_{r=1}^{k}(-1)^{r-1}\frac{B_r}{2r!}\{f^{(2r-1)}(n+1)-f^{(2r-1)}(1)\}+$$
$$+\frac{(-1)^k}{2^{2k-1}\pi^{2k}}\int\limits_1^{n+1} f^{(2k)}(t)\sum_{l=1}^{\infty}\frac{\cos 2\pi lt}{l^{2k}}\,dt.$$

This is equivalent to (13.5.12), with $n+1$ for n and $k-1$ for k.

13.9. A formula of Fourier. We can now give an account of Fourier's formula (2.9.2). This was

(13.9.1) $$\tfrac{1}{2}\pi f(x) = \sum_{h=0}^{\infty}(-1)^h f^{(2h)}(\pi)\left(\sin x-\frac{\sin 2x}{2^{2h+1}}+...\right)$$
$$= \sum_{h=0}^{\infty}\frac{2^{2h}\pi^{2h+1}}{(2h+1)!}f^{(2h)}(\pi)B_{2h+1}\left(\frac{x+\pi}{2\pi}\right),$$

where $f(x)$ is odd and $-\pi < x < \pi$. If we write

$$\frac{x+\pi}{2\pi} = y, \qquad x = 2\pi(y-\tfrac{1}{2}), \qquad f(x) = f\{2\pi(y-\tfrac{1}{2})\} = g(y),$$

so that $0 < y < 1$, it becomes

(13.9.2) $$g(y) = 2\sum_{h=0}^{\infty}\frac{g^{(2h)}(1)}{(2h+1)!}B_{2h+1}(y).$$

Now (13.5.4), if we assume that $\rho_l \to 0$, gives

(13.9.3) $$g(y) = \int\limits_0^1 g(t)\,dt+\sum_{1}^{\infty}\frac{B_\nu(y)}{\nu!}\{g^{(\nu-1)}(1)-g^{(\nu-1)}(0)\}.$$

Also $f(x)$ is odd, so that $g(1-y) = -g(y)$; and hence

$$\int\limits_0^1 g(t)\,dt = 0, \qquad g^{(2h-1)}(1)-g^{(2h-1)}(0) = 0, \qquad g^{(2h)}(1)-g^{(2h)}(0) = 2g^{(2h)}(1),$$

and (13.9.3) reduces to (13.9.2).

Suppose, for example, that $f(x)$ is an integral function of exponential type less than 1, so that $g(y)$ is of type less than 2π. Then $g^{(l)}(y) = O(c^l)$, where $0 < c < 2\pi$, uniformly in (0, 1), and $B_l(y)$ is $O\{(2\pi)^{-l}l!\}$, also uniformly, by (13.3.2) and (13.3.3). Thus $\rho_l \to 0$, and Fourier's formula is valid.

13.10. The case $f(x) = x^{-s}$ and the Riemann zeta-function. We now consider the case in which $f(x) = x^{-s}$ and $a = 1$, so that

(13.10.1)
$$F(x) = F(x,s) = \int_1^x t^{-s}\,dt = \frac{x^{1-s}-1}{1-s} \quad (s \neq 1), \qquad \log x \quad (s = 1),$$

and $F(x,s)$ is an integral function of s.

We suppose first that s is real and $s \geqslant S$, where $S \leqslant 0$. Then (13.6.1) and (13.6.2) are true for $2k > -S-1$. Hence, writing $s^{(p)}$ for $s(s+1)...(s+p)$, we obtain

(13.10.2)
$$\sum_1^n m^{-s} - \frac{n^{1-s}-1}{1-s} - \tfrac{1}{2}n^{-s} + \sum_1^k (-1)^{r-1} s^{(2r-2)} \frac{B_r}{(2r)!} n^{-s-2r+1} \to C(s),$$

(13.10.3)
$$C(s) = \frac{1}{2} + \sum_1^k (-1)^{r-1} s^{(2r-2)} \frac{B_r}{(2r)!} - \frac{s^{(2k+1)}}{(2k+2)!} \int_1^\infty \psi_{2k+2}(t) t^{-s-2k-2}\,dt$$

if $s \neq 1$ and $2k > -S-1$. Also

$$R_{k,n} = \frac{s^{(2k+1)}}{(2k+2)!} \int_n^\infty \psi_{2k+2}(t) t^{-s-2k-2}\,dt = O(n^{-s-2k-1}).$$

Thus

(13.10.4)
$$\sum_1^n m^{-s} - \frac{n^{1-s}-1}{1-s} \sim C(s) + \tfrac{1}{2}n^{-s} - \sum (-1)^{r-1} s^{(2r-2)} \frac{B_r}{(2r)!} n^{-s-2r+1}$$

in the notation of § 2.5 (with n^{-1} for x).

We have supposed $s \neq 1$, but our formulae are still valid, with $F(n) = \log n$, for $s = 1$. In particular

$$1 + \tfrac{1}{2} + ... + \frac{1}{n} - \log n \to C(1),$$

so that $C(1)$ is Euler's constant γ. We thus obtain the formulae

(13.10.5)
$$\gamma = \frac{1}{2} + \sum_1^k (-1)^{r-1} \frac{B_r}{2r} - \int_1^\infty \psi_{2k+2}(t) \frac{dt}{t^{2k+3}}$$

and

(13.10.6)
$$\gamma = \frac{1}{2} + \frac{B_1}{2} - \frac{B_2}{4} + \frac{B_3}{6} - ...,$$

the last series being semi-convergent in the sense of § 13.7.

We now consider complex s. There is then no question of semi-convergence. But if $s = \sigma + i\tau$ then $R_{k,n} = O(n^{-\sigma-2k-1})$, uniformly in τ,

and our other conclusions stand, with σ for s in the appropriate places. Also (13.10.2) holds uniformly in any closed and bounded region D throughout which $\sigma > S$, so that $C(s)$ is an analytic function of s regular in D and therefore, since S may be any negative number, $C(s)$ is an integral function. Finally, when $\sigma > 1$,

$$C(s) = \lim_{n\to\infty}\left(\sum_1^n m^{-s} - \frac{n^{1-s}-1}{1-s}\right) = \sum m^{-s} - \frac{1}{s-1} = \zeta(s) - \frac{1}{s-1}.$$

We have thus proved

THEOREM 245. *Riemann's function $\zeta(s)$ is an analytic function of s, regular all over the plane except for a simple pole at $s = 1$, where it behaves like* $\dfrac{1}{s-1}+\gamma+\dots$.

We have also obtained a series of analytical representations of $\zeta(s)$, such as

(13.10.7) $$\zeta(s) = \lim_{n\to\infty}\left\{\sum_1^n m^{-s} - \frac{n^{1-s}}{1-s} - \tfrac{1}{2}n^{-s}\right\} \quad (\sigma > -1),$$

(13.10.8) $$\zeta(s) = \lim_{n\to\infty}\left\{\sum_1^n m^{-s} - \frac{n^{1-s}}{1-s} - \tfrac{1}{2}n^{-s} + \tfrac{1}{12}sn^{-s-1}\right\} \quad (\sigma > -3),$$

(13.10.9) $$\zeta(s) = \frac{1}{s-1} + \frac{1}{2} - \frac{s(s+1)}{2}\int_1^\infty \frac{\psi_2(t)}{t^{s+2}}\,dt \quad (\sigma > -1),$$

and so on. The formulae require modification when $s = 1$: thus in (13.10.9) we must replace $\zeta(s) - \dfrac{1}{s-1}$ by γ.

When $s = 0$ and $s = -1$, (13.10.7) and (13.10.8) give

$$\sum_1^n 1 - n - \tfrac{1}{2} \to \zeta(0), \qquad \sum_1^n m - \tfrac{1}{2}n^2 - \tfrac{1}{2}n - \tfrac{1}{12} \to \zeta(-1),$$

and show incidentally that

(13.10.10) $$\zeta(0) = -\tfrac{1}{2}, \qquad \zeta(-1) = -\tfrac{1}{12},$$

(13.10.11) $$1+1+1+\dots = -\tfrac{1}{2}, \qquad 1+2+3+\dots = -\tfrac{1}{12} \ (\mathfrak{R}, 0).$$

13.11. The case $f(x) = \log(x+c)$ and Stirling's theorem. We can treat the function $f(x) = x^{-s}\log x$ similarly, and it is plain that the results may be obtained from the corresponding results for x^{-s} by formal differentiation with respect to s. Thus, for example,

(13.11.1) $$\sum_1^n m^{-s}\log m - \frac{n^{1-s}\log n}{1-s} + \frac{n^{1-s}}{(1-s)^2} - \tfrac{1}{2}n^{-s}\log n \to -\zeta'(s)$$

for $\sigma > -1$. If we take $s = 0$ in (13.11.1) we obtain a formula for $\log n!$ which embodies a form of Stirling's theorem; but this case is so important that it is better to treat it independently.

In order to obtain a general formula for $\log \Gamma(x+1)$, not restricted to integral values of x, we take

$$f(x) = \log(x+c) \quad (c > -1), \qquad a = -c,$$

and then write x for $n+c$. We thus obtain

(13.11.2) $$\log \Gamma(x+1) = \sum_1^n \log(m+c) + \log \Gamma(1+c)$$
$$= (x+\tfrac{1}{2})\log x - x + C + S_k(n) + R_{k,n},$$

where

(13.11.3) $$C = \log \Gamma(1+c) - (\tfrac{1}{2}+c)\log(1+c) + 1 + c - S_k(1) - R_{k,1},$$

(13.11.4) $$S_k(n) = \sum_1^k \frac{(-1)^{r-1}B_r}{(2r-1)2r} x^{-2r+1},$$

(13.11.5) $$R_{k,n} = -\frac{1}{2k+2} \int_n^\infty \frac{\psi_{2k+2}(t)}{(t+c)^{2k+2}}\, dt = O(x^{-2k-1}).$$

In particular, taking $k = 0$,

(13.11.6) $$\log \Gamma(x+1) - (x+\tfrac{1}{2})\log x + x \to C,$$

(13.11.7) $$C = \log \Gamma(1+c) - (\tfrac{1}{2}+c)\log(1+c) + 1 + c + \tfrac{1}{2} \int_1^\infty \frac{\psi_2(t)}{(t+c)^2}\, dt.$$

Here C is prima facie a function $C(c)$ of c. It is in fact independent of c, but to prove this naturally demands a little more knowledge of the properties of $\Gamma(x)$ than we have assumed so far in this chapter. It follows, for example, from Gauss's formula

$$\Gamma(1+c) = \lim_{n\to\infty} \frac{n!\, n^c}{(1+c)(2+c)\ldots(n+c)}$$

that $$\log \Gamma(n+1+c) - \log \Gamma(n+1) - c \log n \to 0$$

and so, after (13.11.6),

$$C(c) - C(0) = \lim\{c \log n - (n+c+\tfrac{1}{2})\log(n+c) + (n+\tfrac{1}{2})\log n + c\} = 0.$$

Thus C is independent of c and is defined by (13.11.7) for any c.

There are many ways of evaluating C in finite terms. The most common is by means of 'Wallis's product' for π (a corollary of the product form of $\sin \pi x$). A more natural method here is to use the theory of $\zeta(s)$, since it follows from (13.11.1) that

(13.11.8) $$C = -\zeta'(0).$$

If we put $s = 1+\epsilon$ in Riemann's functional equation (2.2.2), expand both sides in powers of ϵ, and equate the first coefficients, we find that $\zeta(0) = -\frac{1}{2}$, in agreement with (13.10.10), and $\zeta'(0) = -\frac{1}{2}\log 2\pi$; so that

(13.11.9) $$C = \tfrac{1}{2}\log 2\pi,$$

and we obtain the ordinary form of Stirling's theorem.

We can also calculate C from (13.11.7). If, for example, we take $c = 0$, it gives

(13.11.10) $$C = 1+\frac{1}{2}\int\limits_1^\infty \frac{\psi_2(t)}{t^2}\,dt = 1+\frac{1}{2}\sum\int\limits_n^{n+1} \frac{(t-n)(t-n-1)}{t^2}\,dt$$

$$= 1+\sum\left\{1-(n+\tfrac{1}{2})\log\frac{n+1}{n}\right\}.$$

Now $$-1+(n+\tfrac{1}{2})\log\frac{n+1}{n} = 2\int\limits_0^{\frac{1}{2}} \frac{t^2\,dt}{(n+\frac{1}{2})^2-t^2}$$

and, substituting and summing under the integral sign, we obtain

$$C = 1-\int\limits_0^{\frac{1}{2}}\left(\pi t\tan\pi t-\frac{2t^2}{\frac{1}{4}-t^2}\right)dt = \tfrac{1}{2}\log 2\pi.$$

In conclusion we note the formulae

(13.11.11) $$\log 1+\log 2+\ldots = \tfrac{1}{2}\log 2\pi \quad (\Re, 0),$$

(13.11.12) $$\tfrac{1}{2}\log 2\pi = 1-\frac{B_1}{1.2}+\frac{B_2}{3.4}-\frac{B_3}{5.6}+\ldots,$$

the last series being semi-convergent.

13.12. Generalization of the formulae. There is a generalization of the Euler-Maclaurin formula important in the calculus of finite differences. Its formal genesis is as follows. If we write

$$Df(x) = f'(x), \qquad e^{hD}f(x) = f(x)+hf'(x)+\tfrac{1}{2}h^2f''(x)+\ldots = f(x+h),$$

and interpret $D^{-1}f(x)$ as $F(x)$, then

$$f(x+y)+f(x+y+1)+\ldots+f(x+y+n-2)$$

$$= \{e^{yD}+e^{(y+1)D}+\ldots+e^{(y+n-2)D}\}f(x)$$

$$= \frac{e^{(y+n-1)D}-e^{yD}}{e^D-1}f(x) = \frac{e^{yD}}{e^D-1}\{f(x+n-1)-f(x)\}$$

$$= \left\{\frac{1}{D}+(y-\tfrac{1}{2})+\frac{B_2(y)}{2!}D+\ldots\right\}\{f(x+n-1)-f(x)\};$$

and we may write this as $\Phi(n)-\Phi(1)$, where

$$\Phi(n) = F(x+n-1)+(y-\tfrac{1}{2})f(x+n-1)+\frac{B_2(y)}{2!}f'(x+n-1)+\dots.$$

In particular, if we take $x = 1$, we obtain

(13.12.1) $$f(y+1)+f(y+2)+\dots+f(y+n-1) = F(n)+C(y)+(y-\tfrac{1}{2})f(n)+\frac{B_2(y)}{2!}f'(n)+\dots,$$

where

(13.12.2) $$C(y) = -F(1)-(y-\tfrac{1}{2})f(1)-\frac{B_2(y)}{2!}f'(1)-\dots.$$

When $y = 0$, these formulae agree with those of §§ 13.5–6.

If, for example, $f(x) = \log x$ and $a = 0$, we obtain

(13.12.3)

$$\log\Gamma(n+y) = n\log n-n+(y-\tfrac{1}{2})\log n+C+\frac{B_2(y)}{1.2}\frac{1}{n}-\frac{B_3(y)}{2.3}\frac{1}{n^2}+\dots,$$

where

(13.12.4) $$C = \log\Gamma(1+y)+1-\frac{B_2(y)}{1.2}+\frac{B_3(y)}{2.3}-\dots,$$

and comparison of (13.12.3) with Stirling's theorem shows that

$$C = \tfrac{1}{2}\log 2\pi,$$

independently of y. If we take $y = 0$, or differentiate with respect to y and then take $y = 0$, we obtain (13.11.12) and (13.10.6).

All this analysis is formal. We may discuss the formulae by the methods of §§ 13.5–7, or by the complex method developed in § 13.14. It will be observed that we are led to an asymptotic expansion of $\log\Gamma(n+y)$ in powers of n^{-1}, while the argument of § 13.11 leads to one in powers of $(n-1+y)^{-1}$.

13.13. Other formulae for C. There are other formulae for C which are interesting in themselves and will lead us naturally to the analysis of § 13.14.

We observe first that, for $t > 0$,

$$\int_1^\infty \psi_2(w)e^{-tw}\,dt = \frac{2}{t}\int_1^\infty \{\psi_1(w)-\tfrac{1}{2}\}e^{-tw}\,dt = \frac{2}{t}\sum_1^\infty\int_0^1 (u-\tfrac{1}{2})e^{-t(u+n)}\,du,$$

since $\psi_2' = 2\psi_1-1$ and ψ_1 is $w-n$ in $(n, n+1)$. A simple calculation then gives

(13.13.1) $$J(t) = \int_1^\infty \psi_2(w)e^{-tw}\,dw = \frac{2e^{-t}}{t^2}\left(\frac{1}{t}-\frac{1}{2}-\frac{1}{e^t-1}\right).$$

Suppose now that

$$f(x) = \int_0^\infty e^{-xt}\,d\chi(t) \tag{13.13.2}$$

is absolutely convergent for $x > 0$. Then

$$\int_1^\infty \psi_2(w)f''(w)\,dw = \int_1^\infty \psi_2(w)\,dw \int_0^\infty t^2e^{-wt}\,d\chi(t) \tag{13.13.3}$$

$$= \int_0^\infty t^2J(t)\,d\chi(t) = -2\int_0^\infty e^{-t}\left(\frac{1}{e^t-1}-\frac{1}{t}+\frac{1}{2}\right)d\chi(t),$$

by (13.13.1). Also (13.6.1) and (13.6.2) hold for $k \geqslant 0$, so that we may use (13.6.4) with $k = 0$. We thus obtain

$$C = -F(1)+\tfrac{1}{2}f(1)+\int_0^\infty e^{-t}\left(\frac{1}{e^t-1}-\frac{1}{t}+\frac{1}{2}\right)d\chi(t). \tag{13.13.4}$$

For example, we may take

$$d\chi = \frac{t^{s-1}}{\Gamma(s)}\,dt, \qquad f(x) = x^{-s}, \qquad \sigma > 0, \qquad a = 1, \qquad C = \zeta(s)-\frac{1}{s-1}$$

(as in § 13.10), when we obtain

$$\zeta(s) = \frac{1}{s-1}+\frac{1}{2}+\frac{1}{\Gamma(s)}\int_0^\infty e^{-t}\left(\frac{1}{e^t-1}-\frac{1}{t}+\frac{1}{2}\right)t^{s-1}\,dt$$

(a formula actually valid for $\sigma > -1$). When $s = 1$, this gives

$$\gamma = \frac{1}{2}+\int_0^\infty e^{-t}\left(\frac{1}{e^t-1}-\frac{1}{t}+\frac{1}{2}\right)dt = 1+\int_0^\infty e^{-t}\left(\frac{1}{e^t-1}-\frac{1}{t}\right)dt. \tag{13.13.5}$$

The argument leading to (13.13.4) is valid whenever

$$f'(x) = -\int_0^\infty te^{-xt}\,d\chi(t)$$

is absolutely convergent for $x \geqslant 1$, even if (13.13.2) does not hold. Thus the assumption $d\chi = t^{-1}e^{-ct}\,dt$, where $c > -1$, gives $f''(x) = (x+c)^{-2}$ and

$$\int_1^\infty \frac{\psi_2(w)}{(w+c)^2}\,dw = -2\int_0^\infty e^{-(1+c)t}\left(\frac{1}{e^t-1}-\frac{1}{t}+\frac{1}{2}\right)\frac{dt}{t}.$$

Combining this with (13.11.7) and (13.11.9), we find

$$\tfrac{1}{2}\log 2\pi = \log\Gamma(1+c)-(\tfrac{1}{2}+c)\log(1+c)+1+c-$$

$$-\int_0^\infty e^{-(1+c)t}\left(\frac{1}{e^t-1}-\frac{1}{t}+\frac{1}{2}\right)\frac{dt}{t}.$$

In particular, for $c = 0$,

$$\tfrac{1}{2}\log 2\pi = 1 - \int_0^\infty e^{-t}\left(\frac{1}{e^t-1} - \frac{1}{t} + \frac{1}{2}\right)\frac{dt}{t}. \tag{13.13.6}$$

Also, differentiating with respect to c, and replacing c by $c-1$, we obtain

$$\log c - \frac{\Gamma'(1+c)}{\Gamma(1+c)} = \int_0^\infty e^{-ct}\left(\frac{1}{e^t-1} - \frac{1}{t}\right) dt \tag{13.13.7}$$

for $c > 0$ (or complex c with $\Re c > 0$).

There is another set of formulae, of a different type, due to Abel and Plana. Returning to (13.13.2) we observe that $f(z)$ is an analytic function of z regular for $x = \Re z > 0$, and that

$$q(\xi, \eta) = \frac{1}{2i}\{f(\xi+i\eta) - f(\xi-i\eta)\} = -\int e^{-\xi t}\sin \eta t \, d\chi(t) \tag{13.13.8}$$

for $\xi > 0$.† Also, using a familiar formula,

$$\int \frac{q(\xi,\eta)}{e^{2\pi\eta}-1}\, d\eta = -\int e^{-\xi t}\, d\chi \int \frac{\sin \eta t}{e^{2\pi\eta}-1}\, d\eta = -\frac{1}{2}\int e^{-\xi t}\left(\frac{1}{e^t-1} - \frac{1}{t} + \frac{1}{2}\right) d\chi.$$

Taking $\xi = 1$ and using (13.13.4), we find

$$C = -F(1) + \tfrac{1}{2}f(1) - \frac{1}{i}\int \frac{f(1+i\eta) - f(1-i\eta)}{e^{2\pi\eta}-1}\, d\eta. \tag{13.13.9}$$

In particular we have

$$\zeta(s) = \frac{1}{s-1} + \frac{1}{2} - \frac{1}{i}\int \frac{(1+i\eta)^{-s} - (1-i\eta)^{-s}}{e^{2\pi\eta}-1}\, d\eta, \tag{13.13.10}$$

first for $\sigma > 0$ and then, by analytic continuation, for all s;

$$\gamma = \tfrac{1}{2} + 2\int \frac{\eta}{1+\eta^2}\frac{d\eta}{e^{2\pi\eta}-1}, \tag{13.13.11}$$

$$\tfrac{1}{2}\log 2\pi = 1 - 2\int \frac{\arctan \eta}{e^{2\pi\eta}-1}\, d\eta. \tag{13.13.12}$$

The last formula corresponds to the case $f(x) = \log x$, when $f(x)$ is not actually defined by (13.13.2).

In the next section we shall give a proof of (13.13.9) which does not depend on any special integral representation of $f(x)$. It is plain that the truth of the formula must depend upon assumptions about the behaviour of $f(x)$ in the complex plane; and this leads us, in the next section, to investigate the Euler-Maclaurin formula from a quite

† We return here to the convention that integrals without limits are over $(0, \infty)$.

different point of view. It will, however, be useful to make one preliminary remark of a formal character. If we write

$$\frac{1}{i}\{f(1+i\eta)-f(1-i\eta)\} = 2\sum(-1)^{r-1}\frac{f^{(2r-1)}(1)}{(2r-1)!}\eta^{2r-1},$$

insert this expansion in (13.13.9), integrate formally term by term, and observe that

(13.13.13) $$\int \frac{\eta^{2r-1}}{e^{2\pi\eta}-1}\,d\eta = \frac{B_r}{4r},$$

then we are led back to the series

$$C = -F(1)+\tfrac{1}{2}f(1)-\sum\frac{(-1)^{r-1}B_r}{2r!}f^{(2r-1)}(1),$$

thus connecting (13.13.9) with our earlier analysis.

13.14. Investigation of the Euler-Maclaurin formula by complex integration. We suppose now that $f(z)$ is an analytic function of $z = x+iy$, regular for $x \geqslant \xi$, where $\xi < 1$, and that

(13.14.1) $$e^{-2\pi|y|}|f(x+iy)| \to 0,$$

when $|y| \to \infty$, uniformly in any finite interval (ξ, X) of x. We denote the rectangle defined by $x = 1$, $x = n$, and $y = \pm Y$, with semicircular indentations of radius ρ round 1 and n, by $C(\rho)$; the indentations themselves by $I(\rho)$; and define C as the limit of $C(\rho)-I(\rho)$ when $\rho \to 0$, and C_1 and C_2 as the parts of C above and below the real axis.

By Cauchy's theorem

$$\frac{1}{2\pi i}\int_C \pi\cot\pi z f(z)\,dz = \sum_{m=1}^{n}{}' f(m):$$

here the integrals along the vertical sides of C are principal values, and the dash implies that the extreme terms of the sum are affected by a factor $\frac{1}{2}$. Also

$$\int_{C_1}\pi i f(z)\,dz = -\pi i\int_1^n f(x)\,dx, \qquad \int_{C_2}\{-\pi i f(z)\}\,dz = -\pi i\int_1^n f(x)\,dx$$

and so

(13.14.2) $$\sum_1^n f(m) = \int_1^n f(x)\,dx+\tfrac{1}{2}f(1)+\tfrac{1}{2}f(n)+\frac{1}{2i}\int_C \psi(z)f(z)\,dz,$$

where

$$\psi(z) = \cot\pi z+i = \frac{2i}{1-e^{-2\pi iz}}, \qquad \psi(z) = \cot\pi z-i = \frac{2i}{e^{2\pi iz}-1}$$

on C_1 and C_2 respectively.

It follows from (13.14.1) that the integrals along the horizontal sides of C tend to 0 when $Y \to \infty$; so that (13.14.2) reduces to

(13.14.3) $$\sum_{m=1}^{n} f(m) - \int_{1}^{n} f(x)\,dx - \tfrac{1}{2}f(n) - \tfrac{1}{2}f(1) = Q(n) - Q(1),$$

where $$Q(c) = \frac{1}{2i}\int_{c-i\infty}^{c+i\infty} \psi(z)f(z)\,dz,$$

and both $Q(n)$ and $Q(1)$ are principal values, at $z = n$ and $z = 1$ respectively. Also

$$\psi(1+iy) = \frac{2i}{1-e^{2\pi y}} \quad (y > 0), \qquad \frac{2i}{e^{2\pi|y|}-1} \quad (y < 0),$$

and $\psi(n+iy)$ has the same values. Hence, inserting these values in $Q(n)$ and $Q(1)$, and associating together the contributions of positive and negative y, we obtain

(13.14.4) $$Q(1) = 2\int \frac{q(1,y)}{e^{2\pi y}-1}\,dy, \qquad Q(n) = 2\int \frac{q(n,y)}{e^{2\pi y}-1}\,dy,$$

where $q(x,y)$ is defined by (13.13.8). The integrals now converge in the ordinary sense, since $q(x,y)$ is $O(|y|)$ for small y. If f is real for real z, then $q(x,y)$ is the imaginary part of f.

We now expand $q(n,y)$ by Taylor's theorem, in the form

(13.14.5)
$$q(n,y) = yf'(n) - \frac{y^3}{3!}f'''(n) + \ldots + (-1)^{k-1}\frac{y^{2k-1}}{(2k-1)!}f^{(2k-1)}(n) + q_{2k+1}(n,y),$$

substitute in the integral for $Q(n)$, and use (13.13.13). We thus obtain

(13.14.6) $$\sum_{m=1}^{n} f(m) - \int_{1}^{n} f(x)\,dx - \tfrac{1}{2}f(n) - \sum_{r=1}^{k} (-1)^{r-1}\frac{B_r}{(2r)!}f^{(2r-1)}(n)$$
$$= \tfrac{1}{2}f(1) - 2\int \frac{q(1,y)}{e^{2\pi y}-1}\,dy + R_k(n),$$

where $$R_k(n) = 2\int \frac{q_{2k+1}(n,y)}{e^{2\pi y}-1}\,dy.$$

The last series on the left of (13.14.6) is the $S_k(n)$ of § 13.5. If we can show that $R_k(n) \to 0$ when $n \to \infty$, then we shall obtain

$$\sum_{m=1}^{n} f(m) - \int_{1}^{n} f(x)\,dx - \tfrac{1}{2}f(n) - S_k(n) \to \tfrac{1}{2}f(1) - \frac{1}{i}\int \frac{f(1+iy)-f(1-iy)}{e^{2\pi y}-1}\,dy,$$

in agreement with (13.13.9), since $F(1)$ is 0 when $a = 1$. This is the Abel-Plana formula.

In order to prove that $R_k(n) \to 0$, we must naturally impose much severer restrictions on $f(z)$. We suppose that

$$f^{(r)}(z) = O(|z|^{c-r}) \tag{13.14.7}$$

for a fixed c and each r, when $z \to \infty$ in the half-plane $x \geqslant \xi$. We shall also suppose that $2k+1 > c$. Then, using (13.14.5) and the formula

$$g(y) = g(0)+yg'(0)+\ldots+\frac{y^{2k}}{(2k)!}g^{(2k)}(0)+\frac{y^{2k+1}}{(2k)!}\int_0^1 (1-u)^{2k}g^{(2k+1)}(yu)\,du,$$

with $q(n,y)$ for $g(y)$, we find that

$$q_{2k+1}(n,y) = \frac{(-1)^k y^{2k+1}}{2(2k)!}\int_0^1 (1-u)^{2k}\{f^{(2k+1)}(n+iyu)+f^{(2k+1)}(n-iyu)\}\,du.$$

It follows from (13.14.7), since $|n+iy| \geqslant n$ and $2k+1 > c$, that

$$q_{2k+1}(n,y) = O(|y|^{2k+1}n^{c-2k-1})$$

uniformly in y, and so that

$$R_k(n) = O(n^{c-2k-1})\int \frac{y^{2k+1}}{e^{2\pi y}-1}\,dy = O(n^{c-2k-1}) \to 0,$$

when $n \to \infty$; and this completes the proof.

The conditions are satisfied, for example, if $f(x) = x^{-s}$ or $f(x) = \log x$, and we thus recover many of the results of §§ 13.10–13. We naturally cannot expect to find in this way such precise results as those of § 13.7.

13.15. Summability of the Euler-Maclaurin series. We shall call

$$\mathfrak{S}(n) = \tfrac{1}{2}f(n)+\frac{B_1}{2!}f'(n)+0-\frac{B_2}{4!}f'''(n)+0+\ldots = \sum_0^\infty a_r, \tag{13.15.1}$$

where

(13.15.2)

$$a_0 = \tfrac{1}{2}f(n), \quad a_{2r-1} = \frac{(-1)^{r-1}B_r}{2r!}f^{(2r-1)}(n) \quad (r \geqslant 1), \quad a_{2r} = 0 \quad (r \geqslant 1),$$

the 'Euler-Maclaurin series' of $f(n)$. We have seen that it is in certain circumstances an asymptotic series for

$$\Phi(n) = f(1)+f(2)+\ldots+f(n)-\int_a^n f(x)\,dx-C, \tag{13.15.3}$$

and the question remains whether it is summable by any of the methods of earlier chapters. The series usually diverges rapidly, so that a rather drastic method of summation will be needed. We shall show that in certain cases, including those which we have considered particularly in

the preceding sections, the series is summable by the method (B*) of § 8.11.

We suppose first that $0 < \delta < 1$, $s > 0$, and

$$f(z) = O(|z|^{-s}) \tag{13.15.4}$$

uniformly in the half-strip $x \geqslant \delta$. Then

$$f^{(2r-1)}(n) = \frac{(2r-1)!}{2\pi i}\int_C \frac{f(u)}{(u-n)^{2r}}\,du,$$

where C is the line $(\delta+i\infty, \delta-i\infty)$; and hence

$$a(t) = \sum_0^\infty a_k \frac{t^k}{k!} = \tfrac{1}{2}f(n)+\frac{B_1}{1!\,2!}tf'(n)-\frac{B_2}{3!\,4!}t^3f'''(n)+\dots \tag{13.15.5}$$

$$= \frac{1}{2\pi i}\int_C \frac{f(u)}{u-n}\left\{\frac{1}{2}+\frac{B_1}{2!}\frac{t}{u-n}-\frac{B_2}{4!}\left(\frac{t}{u-n}\right)^3+\dots\right\}du$$

$$= \frac{1}{2\pi i}\int_C \frac{f(u)}{n-u}\left(\frac{1}{e^w-1}-\frac{1}{w}\right)du = \frac{1}{2\pi it}\int_C f(u)\left(\frac{w}{e^w-1}-1\right)du,$$

where $w = t/(n-u)$. The integration term by term is justified for small t, since $|w| \leqslant |t|/(n-\delta)$,

$$\frac{B_1}{2!}|w|+\frac{B_2}{4!}|w|^3+\dots = \frac{1}{|w|}-\tfrac{1}{2}\cot\tfrac{1}{2}|w|$$

is bounded for $|w| < \pi$, and

$$\int_C \left|\frac{f(u)}{n-u}\right||du| < \infty.$$

Thus the final formula (13.15.5) is true for small t.

Let us assume provisionally that the function $a(t)$ defined by the series is equal to

$$\frac{1}{2\pi it}\int_C f(u)\left(\frac{w}{e^w-1}-1\right)du$$

for all positive t. Then

$$\int e^{-t}a(t)\,dt = \int \frac{e^{-t}}{t}\left\{\frac{1}{2\pi i}\int_C f(u)\left(\frac{w}{e^w-1}-1\right)du\right\}dt$$

$$= \frac{1}{2\pi i}\int_C f(u)\left\{\int \frac{e^{-t}}{t}\left(\frac{w}{e^w-1}-1\right)dt\right\}du,$$

provided that we may invert the integrations, as we also assume provisionally. The inner integral here is

$$J = \int e^{-t}\left(\frac{\omega}{e^{\omega t}-1}-\frac{1}{t}\right)dt,$$

where $$\omega = \frac{w}{t} = \frac{1}{n-u} = \frac{1}{n-\delta-iy} = \rho e^{i\phi},$$

say, and $|\phi| < \frac{1}{2}\pi$. Applying Cauchy's theorem to the sector bounded by the real axis of t and the radius $\arg t = -\phi$, we obtain

$$J = \int e^{-re^{-i\phi}}\left(\frac{\rho e^{i\phi}}{e^{\rho r}-1}-\frac{e^{i\phi}}{r}\right)e^{-i\phi}\,dr = \int e^{-R/\omega}\left(\frac{1}{e^{R}-1}-\frac{1}{R}\right)dR$$

$$= \int e^{-R(n-u)}\left(\frac{1}{e^{R}-1}-\frac{1}{R}\right)dR = \log(n-u)-\frac{\Gamma'(1+n-u)}{\Gamma(1+n-u)},$$

by (13.13.7), the logarithm having the value which is real when $u = \delta$. Thus

(13.15.6) $$\int e^{-t}a(t)\,dt = \frac{1}{2\pi i}\int_C f(u)\left\{\log(n-u)-\frac{\Gamma'(1+n-u)}{\Gamma(1+n-u)}\right\}du.$$

It follows (apart from the justification of our provisional assumptions) that $\mathfrak{S}(n)$ is summable (B*) to this sum, and that $\mathfrak{S}(n)-\mathfrak{S}(1)$ is summable (B*) to sum

$$\frac{1}{2\pi i}\int_C f(u)\left\{\log\frac{n-u}{1-u}-\frac{\Gamma'(1+n-u)}{\Gamma(1+n-u)}+\frac{\Gamma'(2-u)}{\Gamma(2-u)}\right\}du$$

$$= \frac{1}{2\pi i}\int_C f(u)\log\frac{n-u}{1-u}\,du-\frac{1}{2\pi i}\int_C f(u)\left(\frac{1}{2-u}+\frac{1}{3-u}+\dots+\frac{1}{n-u}\right)du$$

$$= f(2)+f(3)+\dots+f(n)+\frac{1}{2\pi i}\int_C f(u)\log\frac{n-u}{1-u}\,du.$$

Finally,

$$\frac{1}{2\pi i}\int_C f(u)\log\frac{n-u}{1-u}\,du = \frac{1}{2\pi i}\int_{C'} f(u)\{\log(n-u)-\log(1-u)\}\,du,$$

where C' is a lacet formed by the line (δ, n) taken twice in opposite directions, the singularities at $u = 1$ and $u = n$ being avoided in the usual way by semicircles whose radius is made to tend to zero. The

value of $i\Im\{\log(n-u)-\log(1-u)\}$ is 0 on $(\delta, 1)$, $-\pi i$ on $(1,n)$, πi on $(n,1)$, and 0 on $(1,\delta)$, and so

$$\frac{1}{2\pi i}\int\limits_C f(u)\log\frac{n-u}{1-u}\,du = -\int\limits_1^n f(u)\,du,$$

(13.15.7) $$\mathfrak{S}(n)-\mathfrak{S}(1) = f(2)+f(3)+...+f(n)-\int\limits_1^n f(u)\,du,$$

(13.15.8) $$\mathfrak{S}(n) = f(1)+f(2)+...+f(n)-\int\limits_1^n f(u)\,du-C = \Phi(n),$$

the series being summable (B*), and $C = f(1)-\mathfrak{S}(1)$ being the Euler-Maclaurin constant of $f(x)$ for $a = 1$.†

It remains to justify our provisional assumptions. For this it is sufficient to prove (a) that, if $0 < t_0 < t_1$, the integral

$$I = I(t) = \frac{1}{2\pi i}\int\limits_C \frac{f(u)}{n-u}\left(\frac{1}{e^w-1}-\frac{1}{w}\right)du \quad \left(w = \frac{t}{n-u}\right)$$

converges uniformly in some region including the stretch (t_0, t_1) of the real axis in the plane of t; and (b) that the double integral

$$K = \int\limits_C |f(u)||du|\int \frac{e^{-t}}{t}\left|\frac{w}{e^w-1}-1\right|dt = \int\limits_C \frac{|f(u)|}{|n-u|}\,du\int e^{-t}\left|\frac{1}{e^w-1}-\frac{1}{w}\right|dt$$

is convergent. It is plain, first, that the conditions will be satisfied if

(13.15.9) $$\left|\frac{1}{e^w-1}-\frac{1}{w}\right| < H, \qquad \left|\frac{1}{e^w-1}-\frac{1}{w}\right| < H(1+t),$$

respectively, for all relevant values of t and u. For then I and K are majorized by multiples of

(13.15.10) $$\int\limits_C \frac{|f(u)|}{|n-u|}|du|, \quad \int\limits_C \frac{|f(u)|}{|n-u|}|du|\int (1+t)e^{-t}\,dt = 2\int\limits_C \frac{|f(u)|}{|n-u|}|du|,$$

respectively. It is also plainly sufficient to consider the upper half of C.

(a) Suppose that $t = re^{i\theta}$, where

$$0 < r_1 \leqslant r \leqslant r_2, \qquad |\theta| \leqslant \alpha.$$

The first condition (13.15.9) is certainly satisfied if one or other of the conditions

(13.15.11) $$|w| \leqslant \lambda < 2\pi, \qquad p = \Re w \geqslant \xi > 0$$

† The formulae agree with (13.6.3)–(13.6.5) for $a = 1$, $k = \infty$, since then $F(1) = 0$ and $\frac{1}{2}f(n)+S_k(n)$ becomes $\mathfrak{S}(n)$.

is satisfied for some λ and ξ and all t and u in question. If $u = \delta+iy$, where $y = (n-\delta)\tan\phi > 0$, then $0 < \phi < \frac{1}{2}\pi$ and

$$w = \frac{t}{n-\delta-iy} = \frac{re^{i\theta}}{(n-\delta)(1-i\tan\phi)} = \frac{r\cos\phi}{n-\delta}e^{i(\theta+\phi)},$$

$$|w| = \frac{r\cos\phi}{n-\delta}, \qquad p = \frac{r\cos\phi}{n-\delta}\cos(\theta+\phi).$$

If $\phi \geqslant \frac{1}{2}\pi-2\alpha$, $\cos\phi \leqslant \sin 2\alpha$, and

$$|w| \leqslant \frac{r_2}{n-\delta}\sin 2\alpha = \lambda.$$

If $0 < \phi \leqslant \frac{1}{2}\pi-2\alpha$, $\cos(\theta+\phi) \geqslant \sin\alpha$, $\cos\phi \geqslant \sin 2\alpha$, and

$$p \geqslant \frac{r_1}{n-\delta}\sin\alpha\sin 2\alpha = \xi.$$

If we choose α so that $\lambda < 2\pi$, then one or other of (13.15.11) is satisfied for all relevant t and u.

(*b*) In this case t is real and positive, and the second condition (13.15.9) is satisfied if one or other of

(13.15.12) $$|w| \leqslant \lambda < 2\pi, \qquad \Re w \geqslant \xi/t \quad (\xi > 0)$$

is true for the relevant t and u.

Now either (i) $t \leqslant \lambda|n-u|$, where $0 < \lambda < 2\pi$, in which case $|w| \leqslant \lambda$, or (ii) $t > \lambda|n-u|$, in which case

$$\Re w = \Re\left(\frac{t}{n-\delta-iy}\right) = \frac{t(n-\delta)}{|n-u|^2} > \frac{\lambda^2(n-\delta)}{t} = \frac{\xi}{t}.$$

We have thus proved

THEOREM 246. *If $f(z)$ is regular, and $O(|z|^{-s})$, where $s > 0$, in the half plane $x = \Re z \geqslant \delta$, where $\delta < 1$, then the Euler-Maclaurin series of $f(z)$ is summable* (B*), *to the sum* (13.15.3): *in particular*

$$C = -F(1)+\tfrac{1}{2}f(1)-\frac{B_1}{2!}f'(1)+0+\frac{B_2}{4!}f'''(1)+0-\dots \quad (\mathrm{B}^*).$$

13.16. Additional remarks. (1) We have supposed that $f(z)$ satisfies (13.15.4), so that the integrals (13.15.10) are convergent. If we suppose only that $|f(z)| = O(|z|^c)$, for some c, then

$$\int \frac{|f(u)|}{|n-u|^{2r}}|du|$$

will be convergent for $r > \frac{1}{2}(c+1)$, and we shall still be able to prove the summability of

$$\mathfrak{S}^{(r)}(n) = 0+0+\dots+0+\frac{(-1)^{r-1}}{(2r)!}B_r f^{(2r-1)}(n)+0+\frac{(-1)^r}{(2r+2)!}B_{r+1}f^{(2r+1)}(n)+\dots.$$

We can then pass to $\mathfrak{S}(n)$ by adding on a finite number of terms, and the substance of our conclusions will be unaffected. Thus, if $f(u) = u^{-s}$, where $-1 < s < 0$, or $f(u) = \log u$, we can take $r = 1$.

(2) We have used the (B*) method, which involves the notion of the analytic continuation of $a(t)$. If (as actually happens in the most important cases) $a(t)$ is regular for $\Re t > 0$, then the positive axis of t is included in the Borel polygon of $a(t)$, and the series for $a(t)$ is summable (B). In this case we may say that $\mathfrak{S}(n)$ is summable (B²), i.e. by a repeated application of Borel's integral definition.

Suppose, for example, that $f(z)$ satisfies (13.15.4) uniformly in any sector issuing from δ and excluding the negative axis. Then it is not difficult to show, by a modification of the analysis of § 13.15, that $a(t)$ is regular for $\Re t > 0$, so that $\mathfrak{S}(n)$ is summable (B²). We can also combine this remark with the generalization indicated under (1). In particular, the series

$$\frac{1}{2}+\frac{B_1}{2}+0-\frac{B_2}{4}+0+\frac{B_3}{6}-\ldots, \tag{13.16.1}$$

$$1-\frac{B_1}{1.2}+0+\frac{B_2}{3.4}+0-\frac{B_3}{5.6}+\ldots \tag{13.16.2}$$

are summable (B²), to sums γ and $\frac{1}{2}\log 2\pi$ respectively.

It is easy to verify these assertions directly. For example, for (13.16.1),

$$a(t) = \frac{1}{2}+\frac{B_1}{2!}t-\frac{B_2}{4!}t^3+\frac{B_3}{6!}t^5-\ldots = \frac{1}{t}\left(\frac{t}{e^t-1}-1+t\right),$$

the series being convergent for $0 \leqslant t < 2\pi$ and summable (B) for all positive t; and $\int e^{-t}a(t)$ converges to γ, by (13.13.5).

13.17. The $\Re$ definition of the sum of a divergent series. The formulae (13.10.11) give examples of the '$\Re$' summability of divergent series of positive terms. We can use such equations, as did Euler and Ramanujan, to define the sums of series, such as $1-1+1-\ldots$, of the more usual type; but the definitions which result have a narrow range and demand great caution in their application.

Thus it is natural, after (13.10.11), to write

$$2+4+6+\ldots = 2(1+2+3+\ldots) = 2(-\tfrac{1}{12}) = -\tfrac{1}{6}, \tag{13.17.1}$$

$$1+3+5+\ldots = 2+4+6+\ldots-(1+1+1+\ldots) = -\tfrac{1}{6}+\tfrac{1}{2} = \tfrac{1}{3}, \tag{13.17.2}$$

$$1+2+3+4+\ldots = (1+3+\ldots)+(2+4+\ldots) = \tfrac{1}{3}-\tfrac{1}{6} = \tfrac{1}{6}, \tag{13.17.3}$$

$$1-2+3-4+\ldots = (1+3+\ldots)-(2+4+\ldots) = \tfrac{1}{3}+\tfrac{1}{6} = \tfrac{1}{2}. \tag{13.17.4}$$

The last of these equations contradicts (1.2.17); and the sum to be assigned naturally to $1+2+3+\ldots$, by Euler's principle of § 1.3, is either ∞, 'the value of $(1-x)^{-2}$ for $x = 1$', or $-\frac{1}{12}$, the value of $\zeta(s) = \sum n^{-s}$ for $s = -1$.

The $-\frac{1}{6}$ and $\frac{1}{3}$ in (13.17.1) and (13.17.2) are in fact the C of $f(x) = 2x$ and $f(x) = 2x-1$ (with $a = 0$). But the $1+3+\ldots$ and $2+4+\ldots$ in (13.17.3) and (13.17.4) cannot be interpreted similarly. They must be regarded rather as $1+0+3+0+\ldots$ and $0+2+0+4+\ldots$; and then there is no $f(x)$, of a sufficiently regular type, which assumes the appropriate values. Actually, if we wish our results to be consistent, we must interpret $1+0+3+0+\ldots$ and $0+2+0+4+\ldots$ as the values, when $s = -1$, of

$$1^{-s}+3^{-s}+\ldots = (1-2^{-s})\zeta(s), \qquad 2^{-s}+4^{-s}+\ldots = 2^{-s}\zeta(s).$$

These values are $(-1)(-\frac{1}{12}) = \frac{1}{12}$ and $2(-\frac{1}{12}) = -\frac{1}{6}$; and $\frac{1}{12}\pm(-\frac{1}{6})$ is $-\frac{1}{12}$ or $\frac{1}{4}$.

The sum of $\log 1-\log 2+\log 3-\dots$, according to these principles, will be the value of

$$1^{-s}\log 1-2^{-s}\log 2+3^{-s}\log 3-\dots = (2^{1-s}-1)\zeta'(s)-2^{1-s}\log 2\,.\,\zeta(s)$$

for $s = 0$, in agreement with the (A, λ) sum of § 4.7, with $\lambda_n = \log(n+1)$. Thus we obtain

(13.17.5) $$\log 1-\log 2+\log 3-\dots = \zeta'(0)-2\log 2\,.\,\zeta(0) = -\tfrac{1}{2}\log\tfrac{1}{2}\pi.$$

It is easily verified that this is also the (C, 1) sum of the series. For here

$$s_{2p} = \log\frac{1\,.\,3\dots(2p-1)}{2\,.\,4\dots 2p} = \log\frac{2p!}{2^{2p}(p!)^2} = -\tfrac{1}{2}\log p-\tfrac{1}{2}\log\pi+o(1),$$

$$s_{2p-1} = s_{2p}+\log 2p = \tfrac{1}{2}\log p-\tfrac{1}{2}\log\pi+\log 2+o(1),$$

and so
$$\frac{s_0+s_1+\dots+s_n}{n+1} \to -\tfrac{1}{2}\log\tfrac{1}{2}\pi.$$

NOTES ON CHAPTER XIII

§ 13.1. This chapter does not profess to contain a systematic study of the Euler-Maclaurin formula and its generalizations, such as will be found in books on the calculus of finite differences. We concentrate our attention on those aspects of the formula most closely connected with the subject-matter of earlier chapters. We have naturally made considerable use of the principal text-books, in particular

Jordan, *Calculus of finite differences* (Budapest, 1939);
Milne-Thomson, *The calculus of finite differences* (London, 1933);
Nörlund, *Vorlesungen über Differenzrechnung* (Berlin, 1924);
Steffensen, *Interpolation* (Baltimore, 1927);
Whittaker and Robinson, *The calculus of observations* (London, 1924);

and of the shorter accounts in Bromwich, Ford, and Lindelöf. The only part of the chapter with any particular novelty of substance is §§ 13.15–16, on the summability of the series.

We have not attempted to give detailed references for the many special formulae which occur, particularly those connected with the zeta and gamma functions.

§ 13.2. The notations of different writers vary considerably. Our B_n and $\phi_n(x)$ are those used by Bromwich and by Whittaker and Robinson, and our $B_n(x)$ that of Nörlund. But Nörlund writes

$$\frac{t}{e^t-1} = \sum B_n\frac{t^n}{n!}$$

so that his B_{2m+1}, for $m > 0$, is 0, and his B_{2m} is our $(-1)^{m-1}B_m$; and he is followed by Jordan and Milne-Thomson. This notation has the advantage that $B_n(0) = B_n$.

§ 13.5. The formula was found independently by Euler, *Comm. Petropol.* 6 (1732–3, published in 1738), 68–97, Opera (1), 15, 42–72, and Maclaurin, *Treatise of fluxions* (1742), 672. See Cantor's *History*, vol. 3, 663, and *Enzykl. d. Math. Wiss.*, IA 3 (§ 38) and IE (§ 11). The first serious discussion of the remainder was that of Poisson, *Mémoires de l'Institut*, 6 (1823), 571–602, and the first quite rigorous one that of Jacobi, *JM*, 12 (1834), 263–72 (*Werke*, 6, 64–75).

§ 13.6. The nature of Ramanujan's work on divergent series has to be inferred from passages in his letters and note-books.

§ 13.8. Poisson, l.c. under § 13.5.

§ 13.9. Fourier, l.c. under § 2.8.

§§ 13.10–11. A good many of the formulae here and in § 13.13 will be found in Bromwich, Appendix III.

The most direct and elementary method for the analytic continuation of $\zeta(s)$ is that set out in Landau, *Handbuch*, 270–2: the ideas underlying the argument are similar to those used here, but it is arranged inductively. The methods of Riemann, of which Landau also gives an account, are more elegant and more familiar.

For the last method of calculation of C in § 13.11 see Bromwich, *MM*, 36 (1907), 81–5.

§ 13.12. For details see Milne-Thomson, ch. 8, or Nörlund, ch. 3.

§ 13.14. The argument is substantially that of Ford and Lindelöf.

§§ 13.15–16. The main results of these sections seem to be new. There is a paper by Barnes, *QJM*, 35 (1904), 175–88, in which he considers the summability of the series (in the more general form of § 13.12) by methods of the Borel type; but the analysis is unconvincing.

It is easily verified that the series is convergent if $f(x)$ is an integral function of order 1 and type less than 2π.

§ 13.17. For the (C, 1) sum of $\log 1 - \log 2 + \log 3 - \ldots$ see Bromwich (ed. 1), 351.

APPENDIX I

On the evaluation of certain definite integrals by means of divergent series

1. In § 1.2 we gave a number of examples of the use of divergent series in formal calculations, mainly of the values of definite integrals. We show here how these and similar calculations may be justified.

We observe first that all the ordinary theorems concerning the continuity, integration, or differentiation of the sums of convergent series have analogues for any linear method of summation T defined by (3.1.3) or (3.1.4). We state the theorems for the method (3.1.3), and we confine ourselves to the most obvious analogues of classical tests, in which the functions concerned are continuous and the series uniformly convergent. In what follows $\sum a_n(x)$ is summable (T) to $s(x)$, i.e. $s_n(x) = \sum_{m=0}^{n} a_m(x) \to s(x)$ (T).

THEOREM 247. *If* (i) *$a_n(x)$ is continuous in $\langle a,b\rangle$, for each n;*†
(ii) *$t_m(x) = \sum c_{m,n}\, s_n(x)$ is uniformly convergent in $\langle a,b\rangle$, for each m;*
(iii) *$\sum a_n(x)$ is uniformly summable in $\langle a,b\rangle$ to sum $s(x)$;*
then $s(x)$ is continuous in $\langle a,b\rangle$.

For $t_m(x)$ is continuous for each m, by (ii), and the conclusion follows from (iii). If T is row-finite, as, for example, when it is (C,k), then condition (ii) may be omitted.

THEOREM 248. *Under the same conditions*

$$\sum \int_a^b a_n(x)\,dx = \int_a^b s(x)\,dx \quad (T).$$

For the left-hand side is, by definition,

$$\lim_m \sum c_{m,n} \int_a^b s_n(x)\,dx = \lim_m \int_a^b \sum c_{m,n}\, s_n(x)\,dx = \lim_m \int_a^b t_m(x)\,dx,$$

and the conclusion again follows from (iii).

THEOREM 249. *If* (i) *$a_n'(x)$ is continuous in $\langle a,b\rangle$,‡ for each n;*
(ii) *$\sum c_{m,n}\, s_n'(x)$ is, for each m, uniformly convergent in $\langle a,b\rangle$;*
(iii) *$\sum a_n'(x)$ is uniformly summable in $\langle a,b\rangle$;*
(iv) *$\sum a_n(x)$ is summable in $\langle a,b\rangle$, to $s(x)$;*
then $s'(x)$ exists in the interval $a < x < b$, and is continuous, and

$$\sum a_n'(x) = s'(x) \quad (T).$$

† With the usual gloss (assertion of right-hand or left-hand continuity only) at the ends of the interval.

‡ With the gloss corresponding to that on Theorem 247.

This is a trivial corollary of Theorems 247 and 248. For $f(x)$, the sum of $\sum a_n'(x)$, is continuous in $\langle a, b\rangle$, and

$$s(x)-s(a) = \sum \{a_n(x)-a_n(a)\} = \sum \int_a^x a_n'(y)\,dy = \int_a^x \sum a_n'(y)\,dy = \int_a^x f(y)\,dy,$$

all sums being taken in the T sense. It follows that $s'(x) = f(x)$ in the interval $a < x < b$.

2. We pass to the problems (1), (2), (3), (5), and (6) of § 1.2.† If we use the C definitions, then the transformations in (1) and (2) are covered by Theorems 247 and 249, all the differentiated series being uniformly summable (C, l), for sufficiently large l, in appropriate intervals: alternatively, we may use the A definition. The first integration in (3) is covered by Theorem 248.

The argument in (5) needs more consideration. We suppose that $0 < \phi < \pi$. Then

$$\frac{\cos m\theta - \cos m\phi}{\cos\theta - \cos\phi} = 2\sum \frac{\sin n\phi}{\sin\phi}\cos n\theta(\cos m\theta - \cos m\phi) = \sum a_n \quad (\mathrm{C}, 1),$$

the sums being over $(1,\infty)$, and $0 < \theta < \pi$, $\theta \neq \phi$. The series $\sum a_n$ is uniformly summable in $\langle 0, \phi-\epsilon\rangle$ and $\langle \phi+\epsilon, \pi\rangle$, and so

$$\lim_{N\to\infty}\sum_{n=1}^{N}\left(1-\frac{n}{N+1}\right)\left(\int_0^{\pi} - \int_{\phi-\epsilon}^{\phi+\epsilon}\right)a_n\,d\theta = \left(\int_0^{\pi} - \int_{\phi-\epsilon}^{\phi+\epsilon}\right)\left\{\lim_{N\to\infty}\sum_{n=1}^{N}\left(1-\frac{n}{N+1}\right)a_n\right\}d\theta$$

$$= \left(\int_0^{\pi} - \int_{\phi-\epsilon}^{\phi+\epsilon}\right)\frac{\cos m\theta - \cos m\phi}{\cos\theta - \cos\phi}\,d\theta.$$

The right-hand side tends to the integral (1.2.26) when $\epsilon \to 0$, and it is therefore sufficient to prove that

$$\lim_{N\to\infty}\sum_{n=1}^{N}\left(1-\frac{n}{N+1}\right)\int_{\phi-\epsilon}^{\phi+\epsilon} a_n\,d\theta$$

exists for each ϵ and tends to 0 with ϵ. *A fortiori*, it is sufficient to prove that

$$\sum_{n=1}^{\infty}\int_{\phi-\epsilon}^{\phi+\epsilon} a_n\,d\theta$$

† For (4), see the notes on Ch. I.

has these properties. Now it is easily verified that

$$\int_{\phi-\epsilon}^{\phi+\epsilon} a_n\,d\theta = 2\frac{\sin n\phi}{\sin\phi}\left\{\frac{\sin(m-n)\epsilon}{m-n}\cos(m-n)\phi+\frac{\sin(m+n)\epsilon}{m+n}\cos(m+n)\phi-\right.$$
$$\left.-2\cos m\phi\cos n\phi\frac{\sin n\epsilon}{n}\right\},$$

where $\dfrac{\sin(m-n)\epsilon}{m-n}$ is to be interpreted as ϵ when $m=n$; and m is fixed. It is therefore sufficient to show that the series

$$A = \sum_{n=m+1}^{\infty} \sin n\phi\cos(n-m)\phi\frac{\sin(n-m)\epsilon}{n-m},$$

$$B = \sum_{n=-m+1}^{\infty} \sin n\phi\cos(m+n)\phi\frac{\sin(m+n)\epsilon}{m+n},$$

$$C = -2\cos m\phi\sum_{1}^{\infty}\sin n\phi\cos n\phi\frac{\sin n\epsilon}{n}$$

are convergent (as is obvious), and that their sum tends to 0 with ϵ.† But

$$A=\sum_{k=1}^{\infty}\sin(m+k)\phi\cos k\phi\frac{\sin k\epsilon}{k},\qquad B=\sum_{k=1}^{\infty}\sin(k-m)\phi\cos k\phi\frac{\sin k\epsilon}{k},$$

$$A+B = 2\cos m\phi\sum_{1}^{\infty}\sin k\phi\cos k\phi\frac{\sin k\epsilon}{k} = -C,$$

and so $A+B+C=0$.

Passing to (6) of § 1.2, we select the formula (1.2.28). Since

$$\sin 2\theta+\sin 4\theta+\sin 6\theta+\ldots = \tfrac{1}{2}\cot\theta\ \ (\mathrm{C},1)$$

uniformly in $(\epsilon,\frac{1}{2}\pi)$, we have

$$\lim_{N\to\infty}\sum_{n=1}^{N}\left(1-\frac{n}{N+1}\right)\int_{\epsilon}^{\frac{1}{2}\pi}\theta\sin 2n\theta\,d\theta = \tfrac{1}{2}\int_{\epsilon}^{\frac{1}{2}\pi}\theta\cot\theta\,d\theta,$$

and it is enough to prove that

$$\lim_{N\to\infty}\sum_{n=1}^{N}\left(1-\frac{n}{N+1}\right)\int_{0}^{\epsilon}\theta\sin 2n\theta\,d\theta\to 0$$

when $\epsilon\to 0$; *a fortiori*, to prove that

$$\sum_{1}^{\infty}\int_{0}^{\epsilon}\theta\sin 2n\theta\,d\theta\to 0.$$

† We have rejected the terms corresponding to $n=1,2,\ldots,m$ from A, and added those corresponding to $n=0,-1,\ldots\,-m+1$ to B. Obviously these all tend to zero.

But the series here is

$$\sum \frac{1}{4n^2}(\sin 2n\epsilon - 2n\epsilon\cos 2n\epsilon) = \frac{1}{4}\sum \frac{\sin 2n\epsilon}{n^2} + \tfrac{1}{2}\epsilon\log(2\sin\epsilon),$$

and plainly tends to 0.

Many other integrals may be evaluated similarly. We may mention

$$\int_0^\pi \frac{\theta(\pi-\theta)}{\sin\theta}\,d\theta = 8\sum_0^\infty \frac{1}{(2n+1)^3}, \qquad \int_0^\pi \left\{\frac{\theta(\pi-\theta)}{\sin\theta}\right\}^2 d\theta = 6\pi\sum_1^\infty \frac{1}{n^3},$$

$$\int_0^\pi \theta\sec\theta\,d\theta = -4\sum_0^\infty \frac{(-1)^n}{(2n+1)^2}, \qquad \int_0^\pi \frac{\theta\sin\theta}{\cos\phi-\cos\theta}\,d\theta = 2\pi\log(2\cos\tfrac{1}{2}\phi).$$

The last two integrals are principal values, and $0 < \phi < \pi$.

3. The formulae of § 2 may be derived, in much more general forms, from the theory of the 'conjugate series' of Fourier series. If

$$f(\theta) \sim \tfrac{1}{2}a_0 + \sum (a_n\cos n\theta + b_n\sin n\theta) = \tfrac{1}{2}A_0(\theta) + \sum A_n(\theta),$$

then the conjugate series is

$$\sum (b_n\cos n\theta - a_n\sin n\theta) = \sum B_n(\theta).$$

It is familiar that

$$\sum B_n(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\cot\tfrac{1}{2}(t-\theta)\,dt \tag{3.1}$$

under appropriate conditions, the integral being a principal value at $t = \theta$. For example, the series converges to this value if the integral exists and $f(t)$ is of bounded variation in an interval round $t = \theta$. But (3.1) is the result of writing

$$\tfrac{1}{2}\cot\tfrac{1}{2}(t-\theta) = \sin(t-\theta) + \sin 2(t-\theta) + \ldots,$$

and integrating term by term after multiplication by $f(t)$. In particular

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\cot\tfrac{1}{2}t\,dt = \sum b_n.$$

Formulae with divergent integrals

4. We now consider the formulae of § 1.5. There are theorems for divergent integrals corresponding to those of § 1, which we need not state formally. We can verify at once that the integrals (1.5.8) and (1.5.9) are summable (A) to the values stated, and uniformly in any

interval (m_0, m_1) of positive values. They are also summable (C, k) for sufficiently large k, since it is easy to prove that

$$\int x^k e^{-aix}\,dx = \Gamma(k+1)e^{-\frac{1}{2}(k+1)\pi i}a^{-k-1} \quad (C, l)$$

for $k > -1$, $a > 0$, $l > k$. In fact, if we integrate

$$\int \left(1-\frac{z}{X}\right)^l z^k e^{-aiz}\,dz$$

round the rectangle $(0, X, X-i\infty, -i\infty)$, with suitable indentations at 0 and X, we obtain

$$\int_0^X \left(1-\frac{x}{X}\right)^l x^k e^{-aix}\,dx = e^{-\frac{1}{2}(k+1)\pi i}\int \left(1+\frac{iy}{X}\right)^l y^k e^{-ay}\,dy +$$
$$+e^{\frac{1}{2}(l+1)\pi i}e^{-aiX}X^{-l}\int y^l(X-iy)^k e^{-ay}\,dy.$$

The first term on the right tends to

$$e^{-\frac{1}{2}(k+1)\pi i}\int y^k e^{-ay}\,dy = \Gamma(k+1)e^{-\frac{1}{2}(k+1)\pi i}a^{-k-1},$$

and the second to 0, when $X \to \infty$.

We now prove two theorems concerning the formulae (1.5.10) and (1.5.11).

THEOREM 250. *If* (i) $F(x) = \sum a_n x^n$, (ii) $m > 0$, $\mu > -1$,
(iii) $\sum n!\,a_n x^n$ *has a radius of convergence* $R > m^{-1}$,
(iv) $\int e^{-(\tau+mi)x}x^{\mu}F(x)\,dx$ *is convergent for* $\tau > 0$, *then*

$$\int x^{\mu}e^{-mix}F(x)\,dx = \sum \Gamma(n+\mu+1)e^{-\frac{1}{2}(n+\mu+1)\pi i}m^{-n-\mu-1}a_n, \tag{4.1}$$

the integral being an A *integral. We may replace condition* (iii) *by either of the more general conditions*
(iii′) $\sum n!\,|a_n|m^{-n} < \infty$,
(iii″) $\sum n!\,\sqrt{n}\,a_n(im)^{-n}$ *is convergent.*

THEOREM 251. *If* (i) $\phi(x) = O(e^{\epsilon x})$ *for every* $\epsilon > 0$, *so that*

$$\psi(\tau) = \int e^{-\tau x}\phi(x)\,dx$$

is convergent for $\tau > 0$,
(ii) $\psi(\tau)$ *is regular for* $|\tau| < m$,
(iii) $\sum n!\,a_n x^n$ *has a radius of convergence* $R > m^{-1}$,
(iv) $\int e^{-\tau x}\phi(x)F(x)\,dx$ *is convergent for* $\tau > 0$, *then*

$$\int \phi(x)F(x)\,dx = \sum a_n \int x^n\phi(x)\,dx, \tag{4.2}$$

where $F(x) = \sum a_n x^n$, *all the integrals being* A *integrals.*

If $\phi(x) = x^{\mu}e^{-mix}$, then $\psi(\tau) = \Gamma(\mu+1)(\tau+mi)^{-\mu-1}$, and conditions (i) and (ii) are satisfied. Thus Theorem 251 includes the main clause of Theorem 250.

Proof of Theorem 250. The equations

$$(4.3)\quad \chi(\tau) = \int e^{-\tau x}\phi(x)F(x)\,dx = \int e^{-(\tau+mi)x}x^{\mu}F(x)\,dx$$
$$= \sum a_n \int e^{-(\tau+mi)x}x^{n+\mu}\,dx = \sum \Gamma(n+\mu+1)a_n(\tau+mi)^{-n-\mu-1}$$

are certainly true if

$$\int e^{-\tau x}\sum |a_n|x^{n+\mu}\,dx = \sum \Gamma(n+\mu+1)|a_n|\tau^{-n-\mu-1} < \infty,$$

and therefore for $\tau > m$; and the final series in (4.3) is uniformly convergent in any interval $0 \leqslant \tau \leqslant \tau_0$. Hence $\chi(\tau)$ is an analytic function of τ regular for $\tau > 0$, and

$$\chi(\tau) \to \sum \Gamma(n+\mu+1)a_n(mi)^{-n-\mu-1} = \sum \Gamma(n+\mu+1)e^{-\frac{1}{2}(n+\mu+1)\pi i}a_n\,m^{-n-\mu-1}$$

when $\tau \to 0$. This is (4.1). It is plain that the proof is equally valid under condition (iii′).

As regards (iii″), we have

$$\sum \Gamma(n+\mu+1)a_n(\tau+mi)^{-n-\mu-1} = \sum \Gamma(n+\mu+1)a_n(mi)^{-n-\mu-1}z^{n+\mu+1},$$

where
$$z = \frac{mi}{\tau+mi} = x+iy,$$

$$x = \frac{m^2}{\tau^2+m^2}, \qquad 1-x = \frac{\tau^2}{\tau^2+m^2} \sim \frac{\tau^2}{m^2}, \qquad y = \frac{\tau m}{\tau^2+m^2} \sim \frac{\tau}{m}$$

when $\tau \to 0$. Thus $1-x \sim y^2$, and $z \to 1$ along a path having contact of the first order with the unit circle. It is known† that if $f(z) = \sum c_n z^n$ and $\sum \sqrt{n}\,c_n$ is convergent, then $f(z) \to \sum c_n$ when $z \to 1$ along such a path, and our conclusion follows.

Proof of Theorem 251. We have $|\phi(x)| < He^{\epsilon x}$ for any positive ϵ and an appropriate H; and so

$$\int e^{-\tau x}|\phi(x)|\sum |a_n|x^n\,dx \leqslant H\sum |a_n|\int e^{-(\tau-\epsilon)x}x^n\,dx$$
$$= H\sum n!\,|a_n|(\tau-\epsilon)^{-n-1} < \infty$$

if $\tau-\epsilon > R^{-1}$. It follows that

$$(4.4)\quad \int e^{-\tau x}\phi(x)F(x)\,dx = \sum a_n \int e^{-\tau x}\phi(x)x^n\,dx = \sum (-1)^n a_n \psi^{(n)}(\tau)$$

if $\tau > m$.

† See Hardy and Littlewood, *PLMS* (2), 11 (1912), 411–78 (475, Th. 48).

If $\tau > 0$, then $\psi(u)$ is regular inside a circle with centre $u = \tau$ and radius $\sqrt{(m^2+\tau^2)}$; and so, by Cauchy's inequality,

$$|\psi^{(n)}(\tau)| \leqslant \frac{n!M(\eta)}{\{\sqrt{(m^2+\tau^2)}-\eta\}^n}$$

for any $\eta > 0$ and a corresponding $M(\eta)$. Thus

$$\sum |a_n||\psi^{(n)}(\tau)| \leqslant M(\eta) \sum \frac{n!|a_n|}{\{\sqrt{(m^2+\tau^2)}-\eta\}^n}.$$

Since $m > R^{-1}$, we can choose η so that $\sqrt{(m^2+\tau^2)}-\eta > R^{-1}$ throughout any interval $0 \leqslant \tau \leqslant \tau_0$ of τ; and so the series on the right of (4.4.) is uniformly convergent in this interval. It follows that

$$\chi(\tau) = \int e^{-\tau x}\phi(x)F(x)\,dx$$

is regular for $\Re\tau > 0$, and that

$$\int e^{-\tau x}\phi(x)F(x)\,dx \to \sum (-1)^n a_n \psi^{(n)}(0) = \sum a_n \lim_{\tau\to 0} \int e^{-\tau x}\phi(x)x^n\,dx$$

when $\tau \to 0$. This is (4.2).

5. *Examples.* (i) To illustrate Theorem 250, we take

$$F(x) = J_0(x),$$

$$\sum n!\,a_n x^n = \sum \frac{(-1)^k 2k!}{(k!)^2}\left(\frac{x}{2}\right)^{2k} = \sum (-1)^k \frac{1.3...(2k-1)}{2.4...2k} x^{2k}$$

(so that $R = 1$), and $\mu = 0$. We obtain

$$\int J_0(x)\cos mx\,dx = 0+0+0+... = 0,$$

$$\int J_0(x)\sin mx\,dx = \frac{1}{m}+\frac{1}{2}\frac{1}{m^3}+\frac{1.3}{2.4}\frac{1}{m^5}+... = \frac{1}{\sqrt{(m^2-1)}}$$

for $m > 1$. Both equations are false for $m < 1$, the values of the integrals being $(1-m^2)^{-\frac{1}{2}}$ and 0.†

(ii) We may illustrate Theorem 251 by the integral

$$I(c) = \int J_\alpha(x)J_{\alpha+1}(cx)\,dx,$$

where $c > 0$, $\alpha > -1$. We observe first that

(5.1)
$$\int e^{-\tau x}x^{\nu+1}J_\nu(x)\,dx = \frac{2^{\nu+1}\tau\Gamma(\nu+\frac{3}{2})}{\sqrt{\pi}(\tau^2+1)^{\nu+\frac{3}{2}}}$$

for $\tau > 0$, $\nu > -1$;‡ from which it follows that

(5.2)
$$\int x^{\nu+1+2m}J_\nu(x)\,dx = 0 \ \ (A)$$

† Watson, 386.

‡ Ibid., 386.

for $m = 0, 1, 2, \ldots$. On the other hand

(5.3) $$\int x^{\nu-1}J_\nu(x)\,dx = 2^{\nu-1}\Gamma(\nu) \quad (A)$$

for $\nu > 0$.†

If $c < 1$, we take

$$F(x) = x^{-\alpha-1}J_{\alpha+1}(cx) = a_0+a_2x^2+\ldots, \qquad \phi(x) = x^{\alpha+1}J_\alpha(x).$$

Then $R = 1/c > 1$, and $\psi(\tau)$ is regular for $|\tau| < 1$, by (5.1). The conditions of Theorem 251 are satisfied, and

$$I(c) = 0+0+0+\ldots = 0,$$

by (5.2). If, on the other hand, $c > 1$, we take

$$F(x) = x^{-\alpha}J_\alpha(x) = a_0+a_2x^2+\ldots, \qquad \phi(x) = x^\alpha J_{\alpha+1}(cx).$$

Then $R = 1$, and $\quad \psi(\tau) = \int e^{-\tau x}x^\alpha J_{\alpha+1}(cx)\,dx$

is regular for $|\tau| < c$.‡ The conditions of the theorem are again satisfied, and

$$I(c) = a_0\int x^\alpha J_{\alpha+1}(cx)\,dx+0+0+\ldots = \frac{2^{-\alpha}c^{-\alpha-1}}{\Gamma(\alpha+1)}\int x^\alpha J_{\alpha+1}(x)\,dx = c^{-\alpha-1},$$

by (5.3). Thus $I(c) = 0$ if $c < 1$ and $I(c) = c^{-\alpha-1}$ if $c > 1$.§

6. We conclude this appendix by considering some integrals which combine some of the features of those of §§ 2 and 4.

From
$$\tan x = 2(\sin 2x-\sin 4x+\sin 6x-\ldots),$$
$$\sec x = 2(\cos x-\cos 3x+\cos 5x-\ldots),$$

we deduce formally

(6.1) $$\int f(x)\tan x\,dx = 2(v_2-v_4+v_6-\ldots),$$
$$\int f(x)\sec x\,dx = 2(u_1-u_3+u_5-\ldots),$$

where $$u_n = \int f(x)\cos nx\,dx, \qquad v_n = \int f(x)\sin nx\,dx.$$

The integrals on the left in (6.1) will usually be principal values at $\frac{1}{2}\pi$, $\frac{3}{2}\pi,\ldots$ (and may require additional conventions); those on the right may be convergent or divergent. Thus, if $\lambda > 0$,

(6.2) $$\int e^{-\lambda x}\tan x\,dx = 2\left(\frac{2}{2^2+\lambda^2}-\frac{4}{4^2+\lambda^2}+\ldots\right),$$
$$\int e^{-\lambda x}\sec x\,dx = 2\left(\frac{\lambda}{1^2+\lambda^2}-\frac{\lambda}{3^2+\lambda^2}+\ldots\right):$$

† Watson, 391. The integral is convergent if $0 < \nu < \frac{3}{2}$.

‡ $\psi''(\tau)$ is an integral of the type (5.1).

§ Watson, 406. Our theorems do not yield the value of $I(c)$ in the limiting case $c = 1$.

in this case each u_n or v_n is convergent, and the integrals on the left may be defined as limits of integrals over $(0, X)$ when $X \to \infty$ in a way which avoids the poles appropriately. Or again

(6.3) $$\int \cos\lambda x \tan x \, dx = 2\left(\frac{2}{2^2-\lambda^2} - \frac{4}{4^2-\lambda^2} + \dots\right),$$

$$\int \sin\lambda x \tan x \, dx = 0,$$

if $\lambda \neq 2n$; and

(6.4) $$\int \cos\lambda x \sec x \, dx = 0$$

$$\int \sin\lambda x \sec x \, dx = -2\left(\frac{\lambda}{1^2-\lambda^2} - \frac{\lambda}{3^2-\lambda^2} + \dots\right),$$

if $\lambda \neq 2n+1$. In this case the u_n and v_n are A integrals, and the integrals on the left require an A convention in addition to the conventions necessitated by the poles.

We consider these formulae a little more closely. We begin by proving

THEOREM 252. *If $f(x)$ is positive and tends steadily to 0 when $x \to \infty$, $f'(x)$ is continuous, and*

$$\int_0^\infty f(x)\tan x \, dx = \lim_{N\to\infty} \int_0^{N\pi} f(x)\tan x \, dx,$$

where the integral on the right is a principal value at $\frac{1}{2}\pi$, $\frac{3}{2}\pi$,..., is convergent, then

$$\int f(x)\tan x \, dx = 2\left(\int f(x)\sin 2x \, dx - \int f(x)\sin 4x \, dx + \dots\right).$$

We need two preliminary remarks.

(*a*) It follows by partial integration† that

$$\int_0^{N\pi} f(x)\tan x \, dx = \tfrac{1}{2}\int_0^{N\pi} f'(x)\log\cos^2 x \, dx.$$

Since $f'(x) \leqslant 0$, $\log\cos^2 x \leqslant 0$, it is necessary and sufficient for convergence that

$$\int f'(x)\log\cos^2 x \, dx < \infty.$$

We may replace $N\pi$ by any sequence x_N which keeps a distance δ away from the poles of $\tan x$.

(*b*) Since

$$v_{2n} = \int f(x)\sin 2nx \, dx = -\frac{1}{2n}\int f'(x)(1-\cos 2nx)\, dx = O\left(\frac{1}{n}\right),$$

† There is no real difficulty in the partial integration, in spite of the infinities of the integrand. See Hardy, *PLMS* (1), 34 (1902), 17–40 (21).

the series $v_2-v_4+\dots$ will be convergent if it is summable (A). It is therefore sufficient to prove that

$$2(v_2 r^2-v_4 r^4+\dots) \to \tfrac{1}{2}\int f'(x)\log\cos^2 x\,dx$$

(assumed finite) when $r \to 1$. Now

$$2(v_2 r^2-v_4 r^4+\dots) = 2\sum_1^\infty (-1)^{n-1}r^{2n}\int f(x)\sin 2nx\,dx$$

$$= -\sum_1^\infty (-1)^{n-1}\frac{r^{2n}}{n}\int f'(x)(1-\cos 2nx)\,dx,$$

and $\displaystyle\sum \frac{r^{2n}}{n}\int |f'(x)|(1-\cos 2nx)\,dx < 2\sum\frac{r^{2n}}{n}\int |f'(x)|\,dx < \infty$

when $r<1$. Hence

$$2(v_2 r^2-v_4 r^4+\dots) = -\int f'(x)\sum_1^\infty \frac{(-1)^{n-1}r^{2n}}{n}(1-\cos 2nx)\,dx$$

$$= -\frac{1}{2}\int f'(x)\log\frac{(1+r^2)^2}{1+2r^2\cos 2x+r^4}\,dx;$$

and it is sufficient to prove that

$$J(r) = -\int f'(x)\log\frac{(1+r^2)^2}{1+2r^2\cos 2x+r^4}\,dx \to \int f'(x)\log\cos^2 x\,dx$$

when $r \to 1$. But it is easily verified that

$$0 \leqslant \log\frac{(1+r^2)^2}{1+2r^2\cos 2x+r^4} \leqslant \log\frac{1}{\cos^2 x}$$

for $0 \leqslant r \leqslant 1$, $\cos^2 x \neq 0$. It follows that $J(r)$ converges uniformly for $0 \leqslant r \leqslant 1$, and that $J(r) \to J(1)$. This proves Theorem 252. The conditions are satisfied if $f(x) = e^{-\lambda x}$, when we obtain the first equation (6.2).

The proof of the second formula is similar.

To prove (6.3) we observe first that

$$\int e^{-sx}\log\cos^2 x\,dx$$

converges, and represents a regular function of s, for $\Re s > 0$. The same is true of

$$\int e^{-sx}\tan x\,dx = -\tfrac{1}{2}s\int e^{-sx}\log\cos^2 x\,dx,$$

where the integral on the left is defined as under (a). Putting $s=\sigma+i\lambda$, and making $\sigma \to 0$, we obtain (6.3). The formulae (6.4) may be proved similarly.

APPENDIX II

The Fourier kernels of certain methods of summation

1. It is familiar that the summability of a Fourier series

$$\tfrac{1}{2}a_0+\sum(a_n\cos n\theta+b_n\sin n\theta)$$

by a method T with

$$\tau_m=\sum c_{m,n}s_n\quad [\textit{or}\ \ \tau(x)=\sum c_n(x)s_n]\dagger$$

depends on the properties of the 'kernel'

$$K_m(t)=\sum c_{m,n}\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}\quad\left[\textit{or}\ \ K(x,t)=\sum c_n(x)\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t}\right].$$

In particular, if T is a 'regular K-method', in the sense of Hardy and Rogosinski,‡ i.e. if it is regular and

$$\sum n|c_{m,n}|<\infty\quad[\textit{or}\ \ \sum n|c_n(x)|<\infty]$$

for each m [or x], then a necessary and sufficient condition for summability, to sum c, is that

$$\int_0^\delta g_c(t)K_m(t)\,dt\to 0\quad\left[\textit{or}\ \int_0^\delta g_c(t)K(x,t)\,dt\to 0\right],$$

where
$$g_c(t)=\tfrac{1}{2}\{f(\theta+t)+f(\theta-t)-2c\},$$
for any $\delta>0$.§

Further, if we call the conditions

(1.1) $\quad g_c(t)\to 0,\qquad$ (1.2) $\quad \displaystyle\int_0^t g_c(u)\,du=o(t),\qquad$ (1.3) $\quad\displaystyle\int_0^t |g_c(u)|\,du=o(t)$

k_c, l_c, and L_c, respectively, then there are three fundamental theorems concerning regular K-methods, due in essentials to Lebesgue.‖

A. If

(1.4)
$$\int_0^\pi |K_m(t)|\,dt<H,$$

where H is independent of m, then k_c is a sufficient condition for summability to c.

† We use τ for the t of § 3.1, t being needed for other purposes.
‡ *Fourier series*, ch. 5 (referred to as *HR* in what follows).
§ *HR*, Theorem 69.
‖ See *HR*, Theorems 70, 71, 72. In Theorem 72, the first condition (5.6.5) is a consequence of the second, provided that $K_m^*(\pi)=O(1)$. The variations for the continuous parameter x are trivial.

B. If

$$\int_0^{\pi} t|K'_m(t)|\,dt < H, \tag{1.5}$$

then l_c is a sufficient condition.

C. If $|K_m(t)| \leqslant K^*_m(t)$, $K^*_m(t)$ is absolutely continuous, except perhaps at the origin, and

$$\int_0^{\pi} t|K^{*\prime}_m(t)|\,dt < H, \qquad |K^*_m(\pi)| < H, \tag{1.6}$$

then L_c is a sufficient condition.

It is plain that k_c implies L_c and that L_c implies l_c; on the other hand, (1.5) implies (1.6) and (1.6) implies (1.4). The condition L_c, and *a fortiori* l_c, is satisfied, with $c = f(\theta)$, for almost all θ; so that a method of summation which satisfies (1.6), and *a fortiori* one which satisfies (1.5), is 'Fourier-effective', i.e. sums any Fourier series almost everywhere to its generating function.

When T is (C, 0), (C, 1), A, then $K_m(t)$, or $K(r,t)$,† is

$$D_m(t) = \frac{\sin(m+\frac{1}{2})t}{2\sin\frac{1}{2}t}, \qquad F_m(t) = \frac{1}{2(m+1)}\left\{\frac{\sin\frac{1}{2}(m+1)t}{\sin\frac{1}{2}t}\right\}^2,$$

$$P(r,t) = \frac{1-r^2}{2(1-2r\cos t+r^2)}.$$

The last kernel satisfies (1.5); the second satisfies (1.6), and *a fortiori* (1.4), but not (1.5); while the first does not satisfy even (1.4). Thus the A and (C, 1) methods are Fourier-effective. On the other hand, classical convergence is not.‡

We consider here the kernels of the methods (C, k), (A, 2), (V P), B, (E, q).

2. The (C, k) kernel. We prove

THEOREM 253. *The* (C, k) *kernel satisfies the conditions of C, and a* fortiori *that of A, for every positive k. It satisfies that of B if $k > 1$, but not if $0 < k \leqslant 1$.*

It is a corollary that the (C, k) method is Fourier-effective for every positive k. We write

$$\tfrac{1}{2}+\cos t+\cos 2t+\ldots = c_0+c_1+c_2+\ldots = \sum c_n,$$

† It is convenient to use r here instead of x.

‡ *HR*, 70 (Theorem 79).

and denote by C_n^k the kth sum-function of $\sum c_n$.† Then

$$K_m(t) = \binom{m+k}{k}^{-1} C_m^k.$$

Plainly we may suppose $m > 2$.

(1) We estimate $K_m(t)$ for $0 < t \leqslant 1/m$ and $1/m < t \leqslant \pi$. If $0 < t \leqslant 1/m$, then $c_m = O(1)$, $C_m = O(m)$, $C_m^k = O(m^{k+1})$, and $K_m = O(m)$, uniformly in t; so that, for an appropriate H,

(2.1) $$|K_m| \leqslant K_m^* = Hm \quad (0 < t \leqslant 1/m).$$

When $t > 1/m$, we use the formulae

$$\sum c_m u^m = \frac{1}{2}\frac{1-u^2}{1-2u\cos t+u^2},$$

$$\sum C_m^k u^m = \frac{1}{2}\frac{1-u^2}{1-2u\cos t+u^2}\frac{1}{(1-u)^{k+1}},$$

(2.2) $$C_m^k = \frac{1}{4\pi i}\int_C \frac{1-u^2}{1-2u\cos t+u^2}\frac{du}{(1-u)^{k+1}u^{m+1}},$$

where C is a small circle round the origin. We may deform C into a lacet C_1 formed by the circle $|u-1| = \rho$, where $\rho < |1-e^{it}|$, and the line $(1+\rho, \infty)$ described twice in opposite directions, provided we allow for the residues at the poles $u = e^{\pm it}$. Calculating the residues, we find

(2.3) $$K_m(t) = \Omega(m) + W(m),$$

(2.4) $$\Omega(m) = \frac{\Gamma(k+1)\Gamma(m+1)}{\Gamma(m+k+1)}\frac{\sin\{(m+\frac{1}{2}k+\frac{1}{2})t-\frac{1}{2}k\pi\}}{(2\sin\frac{1}{2}t)^{k+1}},$$

(2.5) $$W(m) = \frac{\Gamma(k+1)\Gamma(m+1)}{\Gamma(m+k+1)}\frac{1}{4\pi i}\int_{C_1}\frac{1-u^2}{1-2u\cos t+u^2}\frac{du}{(1-u)^{k+1}u^{m+1}}.$$

It is plain that

(2.6) $$\Omega(m) = O(m^{-k}t^{-k-1})$$

uniformly, and we have to estimate $W(m)$.

We take $\rho = \frac{1}{2}m^{-1}$. Then, for a positive H,

$$|1-ue^{\pm it}| = |u-e^{\mp it}| > Ht$$

(since $t > m^{-1}$) on the circular part of C_1, and *a fortiori* on the rectilinear parts. Also $|u|^{-m-1}$ is bounded on the circular part. Hence the circular part contributes

$$O\left(m^{-k}\cdot\frac{m^{-1}}{t^2}\cdot\frac{m^{-1}}{m^{-k-1}}\right) = O\left(\frac{1}{mt^2}\right).$$

† C_n^k corresponds to the A_n^k of Ch. V.

and the rectilinear parts contribute

$$O\left(m^{-k}\cdot\frac{1}{t^2m^{-k}}\int\limits_{1+\frac{1}{2}m^{-1}}^{\infty}\frac{du}{u^m}\right)=O\left\{\frac{1}{t^2(m-1)}\left(1+\frac{1}{2m}\right)^{-m+1}\right\}=O\left(\frac{1}{mt^2}\right).$$

It now follows from (2.3)–(2.6) that

$$K_m(t)=O(m^{-k}t^{-k-1})+O(m^{-1}t^{-2}), \tag{2.7}$$

$$|K_m(t)|\leqslant K_m^*(t)=\tfrac{1}{2}H(m^{-k}t^{-k-1}+m^{-1}t^{-2})\quad (m^{-1}<t\leqslant\pi), \tag{2.8}$$

with an H which we may suppose the same as in (2.1). The K_m^* defined by (2.1) and (2.8) is absolutely continuous, $K_m^*(\pi)=O(1)$, and

$$\int\limits_0^{\pi}t|K_m^{*\prime}|\,dt=\tfrac{1}{2}H\left\{(k+1)m^{-k}\int\limits_{1/m}^{\pi}t^{-k-1}\,dt+2m^{-1}\int\limits_{1/m}^{\pi}t^{-2}\,dt\right\}\leqslant\frac{3k+1}{2k}H.$$

This proves the first clause of Theorem 253.

(2) We now suppose $k>1$ and estimate $K_m'(t)$. Since K_m' is derived from $0-\sin t-2\sin 2t-\dots$ as K_m is from $\frac{1}{2}+\cos t+\cos 2t+\dots$, we have $K_m'(t)=O(m^2)$ and

$$\int\limits_0^{1/m}t|K_m'(t)|\,dt=O\left(m^2\int\limits_0^{1/m}t\,dt\right)=O(1). \tag{2.9}$$

When $t>1/m$, $K_m'=\Omega'+W'$, where Ω' and W' are the derivatives of Ω and W with respect to t. Thus, first,

$$\Omega'(m)=O(m^{-k}.m.t^{-k-1})+O(m^{-k}t^{-k-2})=O(m^{1-k}t^{-k-1}).$$

Next,

$$W'(m)=-\frac{\Gamma(k+1)\Gamma(m+1)}{\Gamma(m+k+1)}\frac{1}{2\pi i}\int\frac{(1-u^2)u\sin t}{(1-2u\cos t+u^2)^2}\frac{du}{(1-u)^{k+1}u^{m+1}}.$$

Hence, if we treat $W'(m)$ as we treated $W(m)$ under (1), we obtain two terms, of which the first is

$$O\left(m^{-k}.\frac{m^{-1}t}{t^4}\cdot\frac{m^{-1}}{m^{-k-1}}\right)=O\left(\frac{1}{mt^3}\right),$$

and the second is

$$O\left(m^{-k}\cdot\frac{t}{t^4m^{-k}}\int\limits_{1+\frac{1}{2}m^{-1}}^{\infty}\frac{du}{u^{m-1}}\right)=O\left\{\frac{1}{t^3(m-2)}\left(1+\frac{1}{2m}\right)^{-m+2}\right\}=O\left(\frac{1}{mt^3}\right).$$

It follows that

$$W'(m)=O(m^{-1}t^{-3}),\qquad K_m'(t)=O(m^{1-k}t^{-k-1})+O(m^{-1}t^{-3}),$$

$$\int\limits_{1/m}^{\pi}t|K_m'|\,dt=O\left(m^{1-k}\int\limits_{1/m}^{\pi}t^{-k}\,dt\right)+O\left(m^{-1}\int\limits_{1/m}^{\pi}t^{-2}\,dt\right)=O(1), \tag{2.10}$$

since $k>1$. This, with (2.9), shows that K_m satisfies (1.5).

The argument fails when $k \leqslant 1$, since then the first term in (2.10) is not $O(1)$, and in fact the (C, 1) method does not satisfy (1.5).† When $k < 0$, the method is not regular, and (since summability then implies convergence) is not Fourier-effective. Kogbetliantz [*AEN* (3), 40 (1923), 259–323 (276)] has proved that *if* $-1 < k \leqslant 1$ *then*

$$K_m(t) = O(m)$$

and

$$K_m(t) = \Omega(m) + O(m^{-1}t^{-2})$$

uniformly for $0 < t \leqslant \pi$. The proof may be carried out on the lines of that of Theorem 253.

3. The (A, 2) and (V P) kernels. For the (A, 2) method

$$K(r,t) = \tfrac{1}{2} + r\cos t + r^4 \cos 2t + r^9 \cos 3t + \dots.$$

If

$$r = e^{-\eta^2\pi^2} \quad (\eta > 0)$$

then, for $0 < t < \pi$,

$$K = \tfrac{1}{2} + \sum_1^\infty e^{-(n\eta\pi)^2} \cos nt = \frac{1}{2\eta\sqrt{\pi}} \sum_{-\infty}^{\infty} e^{-(2m\pi-t)^2/4\eta^2\pi^2} \tag{3.1}$$

$$= \frac{1}{2\eta\sqrt{\pi}} e^{-t^2/4\pi^2\eta^2} + O(e^{-H/\eta^2}),$$

uniformly in t. For summability to c, it is necessary and sufficient that

$$\frac{1}{\eta} \int_0^\delta g_c(t) e^{-t^2/4\pi^2\eta^2}\, dt \to 0$$

when $\eta \to 0$. The kernel K mimics

$$L = \frac{1}{2\eta\sqrt{\pi}} e^{-t^2/4\pi^2\eta^2},$$

and its derivatives mimic those of L. Also

$$\int_0^\pi t|L'(t)|\, dt = O\left(\frac{1}{\eta}\int_0^\infty t \cdot \frac{t}{\eta^2} \cdot e^{-t^2/4\pi^2\eta^2}\, dt\right) = O\left(\eta^{-3}\int_0^\infty t^2 e^{-t^2/4\pi^2\eta^2}\, dt\right) = O(1).$$

Thus the (A, 2) method satisfies (1.5). It is easily verified that it satisfies

$$\int_0^\pi t^p |K^{(p)}(t)|\, dt = O(1) \tag{3.2}$$

for every p. It follows that, like the A method, it will sum derived series of Fourier series at points where the generating function has a derivative, or generalized derivative, of appropriate order.‡

† See *HR*, 62.

‡ See *HR*, 68–9, or Zygmund, ch. 10 (where there is a much fuller account).

We shall prove in Appendix V that the (A, k) method is Fourier-effective for every positive k.

It is interesting to find a row-finite method which has similar properties, and de la Vallée-Poussin's provides an example. In this case

$$K_m(t) = \frac{1}{2}+\frac{m}{m+1}\cos t+\frac{m(m-1)}{(m+1)(m+2)}\cos 2t+\dots = A_m(\cos\tfrac{1}{2}t)^{2m},$$

where
$$A_m = 2^{2m-1}\frac{(m!)^2}{(2m)!} \sim \tfrac{1}{2}\sqrt{(m\pi)}.$$

Since $\cos x < e^{-\frac{1}{2}x^2}$ for $0 < x \leqslant \frac{1}{2}\pi$,

$$\int_0^{\pi} t|K_m'(t)|\,dt = O\left\{m^{\frac{3}{2}}\int_0^{\pi} t\sin\tfrac{1}{2}t(\cos\tfrac{1}{2}t)^{2m-1}\,dt\right\}$$
$$= O\left(m^{\frac{3}{2}}\int_0^{\infty} t^2 e^{-\frac{1}{4}mt^2}\,dt\right) = O(1).$$

Thus $K_m(t)$ satisfies (1.5), and it may be verified that it satisfies (3.2) for general p.

4. The B and E kernels. Since a_n and b_n tend to 0, it is immaterial which form of Borel's definition we choose. Taking the exponential definition, we have

$$K(x,t) = e^{-x}\sum\frac{x^n}{n!}\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t} = \frac{e^{-x(1-\cos t)}}{2\sin\frac{1}{2}t}\sin(x\sin t+\tfrac{1}{2}t);$$

and we may replace this, with trivial error, by

$$t^{-1}e^{-x(1-\cos t)}\sin(x\sin t).$$

The kernel does not satisfy (1.4). For $e^{-x(1-\cos t)} > H$ for $0 < t < x^{-\frac{1}{2}}$ and

$$|\sin(x\sin t)| > H|\sin xt| + O(xt^3).$$

It follows that

$$\int_0^{x^{-\frac{1}{2}}} |K|\,dt > H\int_0^{x^{-\frac{1}{2}}}\frac{|\sin xt|}{t}\,dt + O(x^{-\frac{1}{2}}) > H\int_0^{x^{\frac{1}{2}}}\frac{|\sin u|}{u}\,du \to \infty$$

when $x \to \infty$. In fact the method is not Fourier-effective, and does not sum all Fourier series at points of continuity.

The (E, q) kernel behaves similarly. We suppose $q = 1$, since it is only in this case that the formula for $K_m(t)$ is simple. Then

$$K_m(t) = 2^{-m}\sum_{n=0}^{m}\binom{m}{n}\frac{\sin(n+\frac{1}{2})t}{2\sin\frac{1}{2}t} = \frac{(\cos\frac{1}{2}t)^m}{2\sin\frac{1}{2}t}\sin\frac{m+1}{2}t.$$

It is easily verified that $\int|K_m(t)|\,dt$ is not bounded.

APPENDIX III

On Riemann and Abel summability

1. We prove here three theorems which we have referred to in Chs. IV and XII, and which may be stated shortly as follows.

THEOREM 254: (R, 2) $\rightarrow$ (A).

THEOREM 255: (R, 1) $\rightarrow$ (R, 2).

THEOREM 256: (R_2) $\rightarrow$ (A).

They assert relations of complete inclusion between methods of summation: thus Theorem 254 says that, if a series is summable (R, 2), then it is summable (A) to the same sum.

There are different proofs of all the theorems. We follow that of Kuttner, who first proved Theorem 254 in full generality. The method of proof, by 'formal multiplication' of trigonometrical series, was devised by Rajchmann and developed by Zygmund. It is unlike any method which we have used so far, but depends on theorems which we proved in Ch. X.

We defined the 'Laurent product' of two series, infinite in both directions, in Ch. X. Here we are concerned with trigonometrical series. We write

$$A = \sum a_m e^{mix}, \qquad B = \sum b_n e^{nix},$$

and define C by

$$C = \sum c_p e^{pix}, \qquad c_p = \sum_{m+n=p} a_m b_n.$$

If a_m and b_n are even, and

$$a_m + a_{-m} = 2a_m = \alpha_m, \qquad b_n + b_{-n} = 2b_n = \beta_n,$$

then c_p also is even, $c_p + c_{-p} = 2c_p = \gamma_p$, say, and

$$(\tfrac{1}{2}\alpha_0 + \alpha_1 \cos x + \dots)(\tfrac{1}{2}\beta_0 + \beta_1 \cos x + \dots) = \tfrac{1}{2}\gamma_0 + \gamma_1 \cos x + \dots,$$

where

$$\gamma_p = \tfrac{1}{2} \sum_{m+n=p} \alpha_m \beta_n.$$

This is the formal rule for the multiplication of cosine series.† In our applications, the sums which define the γ_p will all be absolutely convergent. The product of two sine series is also a cosine series, and that of a cosine series and a sine series is a sine series: it will not be necessary to write the formulae down in detail.

† The formulae agree with those for 'Fourier multiplication' in § 10.12: in particular the formula for γ_p agrees with (10.12.8)–(10.12.10) if we remember that α_m and β_n are even.

Theorem 193 shows that, if $a_n = o(1)$, $\sum |n||b_n| < \infty$, and

$$\sum b_n e^{nix} = B(x),$$

then

$$\sum_{-N}^{N} c_n e^{nix} - B(x) \sum_{-N}^{N} a_n e^{nix} \to 0$$

uniformly in x.† Translating this into the language of cosine series, we obtain

THEOREM 257. *If*

$$\alpha_n = o(1), \quad \sum n|\beta_n| < \infty, \quad \tfrac{1}{2}\beta_0 + \sum \beta_n \cos nx = B(x),$$

and

$$\gamma_p = \tfrac{1}{2} \sum_{m+n=p} \alpha_m \beta_n = \tfrac{1}{2}(\alpha_0 \beta_p + \alpha_1 \beta_{p-1} + \ldots + \alpha_p \beta_0) +$$
$$+ \tfrac{1}{2}(\alpha_{p+1}\beta_1 + \alpha_{p+2}\beta_2 + \ldots + \alpha_1 \beta_{p+1} + \alpha_2 \beta_{p+2} + \ldots),‡$$

then

$$\tfrac{1}{2}\gamma_0 + \sum_{1}^{N} \gamma_n \cos nx - B(x)\Big(\tfrac{1}{2}\alpha_0 + \sum_{1}^{N} \alpha_n \cos nx\Big) \to 0$$

uniformly in x.

There are, of course, corresponding theorems for a cosine and a sine series or for two sine series.

2. We require a preliminary lemma.

THEOREM 258. *If* $\sum c_n(1-\cos nx)$ *is convergent for all* x *of an interval* (α, β), *then* $\sum c_n$ *is convergent.*

We use the formula

$$(2.1) \quad \tfrac{1}{2}\pi \int_{\alpha}^{\beta} (1-\cos nx) \sin \pi \frac{x-\alpha}{\beta-\alpha}\, dx = \beta - \alpha + (\cos n\alpha + \cos n\beta) Q_n,$$

where

$$(2.2) \quad Q_n = \frac{\pi H}{2(n^2 - H^2)}, \qquad H = \frac{\pi}{\beta - \alpha}.$$

Plainly,

$$(2.3) \quad 0 < Q_n < \pi H n^{-2} \quad (n^2 > 2H^2).$$

Since $c_n \sin^2 \frac{1}{2}nx \to 0$ in a set of positive measure, $c_n \to 0$ and $\sum n^{-2}|c_n| < \infty$.§

We suppose $\sum c_n$ divergent and deduce a contradiction. If $\sum c_n$ is divergent, then there is a positive δ such that

$$(2.4) \quad \Big|\sum_{\mu}^{\nu} c_n\Big| > \delta$$

† See the remark concerning uniformity on p. 235.

‡ Remembering that $\alpha_{-m} = \alpha_m$, $\beta_{-n} = \beta_n$.

§ See, e.g., *HR*, 84.

for pairs μ, ν tending to infinity. We can therefore find μ_1 and ν_1 so that

$$(2.5)\qquad \mu_1^2 > 2H^2,\quad \nu_1 > \mu_1,\quad \left|\sum_{\mu_1}^{\nu_1} c_n\right| > \delta,\quad \sum_{\mu_1}^{\nu_1} \frac{|c_n|}{n^2} < \frac{(\beta-\alpha)\delta}{4\pi H};$$

and then

$$\left|\tfrac{1}{2}\pi\int_\alpha^\beta \sum_{\mu_1}^{\nu_1} c_n(1-\cos nx)\sin\pi\frac{x-\alpha}{\beta-\alpha}\,dx\right| = \left|\sum_{\mu_1}^{\nu_1} c_n\{\beta-\alpha+(\cos n\alpha+\cos n\beta)Q_n\}\right|$$

$$\geqslant (\beta-\alpha)\left|\sum_{\mu_1}^{\nu_1} c_n\right| - 2\sum_{\mu_1}^{\nu_1}|c_n|Q_n \geqslant (\beta-\alpha)\delta - 2\pi H\sum_{\mu_1}^{\nu_1}\frac{|c_n|}{n^2} > \tfrac{1}{2}(\beta-\alpha)\delta,$$

by (2.1)–(2.5). It follows that

$$(2.6)\qquad \left|\sum_{\mu_1}^{\nu_1} c_n(1-\cos nx)\right| > \sin\pi\frac{x-\alpha}{\beta-\alpha}\left|\sum_{\mu_1}^{\nu_1} c_n(1-\cos nx)\right| > \frac{\delta}{\pi}$$

at a point of (α,β), and so throughout an interval (α_1,β_1) interior to (α,β).

We can now choose a second pair μ_2, ν_2, with $\mu_2 > \nu_1$, for which (2.4) and (2.5) are true, and deduce that (2.6), with μ_1, ν_1, α, β replaced by μ_2, ν_2, α_1, β_1, is true throughout an (α_2,β_2) interior to (α_1,β_1); and we can repeat the argument indefinitely. We thus determine a sequence of pairs μ_k, ν_k, tending to infinity with k, and a corresponding sequence of intervals (α_k,β_k), each included in its predecessor, such that

$$(2.7)\qquad \left|\sum_{\mu_k}^{\nu_k} c_n(1-\cos nx)\right| > \frac{\delta}{\pi}$$

throughout (α_k,β_k). There is at least one x common to all these intervals, and our series diverges for this x; a contradiction which proves the theorem.

3. We now prove Theorem 254. This is the most important of our theorems, and we write out the proof in full, then indicating shortly the points of difference in the proofs of Theorems 255 and 256.

We are given that $\sum a_n = s$ (R, 2), and we may suppose (altering two terms of the series if necessary) that $a_0 = 0$, $s = 0$. Then

$$F(h) = \sum n^{-2}a_n\sin^2\tfrac{1}{2}nh$$

converges for small h, and $F(h) = o(h^2)$; and we have to prove that $\sum a_n r^n \to 0$ when $r \to 1$. The proof falls into two parts: we prove first that the truth of the theorem in two particular cases involves its truth in general, and then prove the particular cases. It is the first stage of the proof which depends on Theorem 257.

Since $\sum n^{-2}a_n$ is convergent, by Theorem 258, we may write our series as

$$\tfrac{1}{2}\alpha_0+\sum \alpha_n \cos nh, \tag{3.1}$$

where

$$\alpha_n = -\tfrac{1}{2}n^{-2}a_n \quad (n>0), \qquad \alpha_0 = -2\sum \alpha_n = \sum n^{-2}a_n.\dagger \tag{3.2}$$

This series converges to $F(h)$ for small h. The two particular cases considered in the proof are those in which (A) the series (3.1) is a Fourier series, (B) the series (3.1) converges uniformly to 0 in an interval of h including the origin.

4. We begin by proving that the theorem, if true in cases (A) and (B), is true generally.

We suppose that (3.1) is convergent and its sum bounded for $|h| \leqslant \delta$, choose a positive η less than $\tfrac{1}{2}\delta$, and suppose that $\lambda(h)$ is any function satisfying the conditions (i) $\lambda(h)$ is even and periodic, and has three continuous derivatives, (ii) $\lambda(h)=1$ for $|h|\leqslant \eta$, (iii) $\lambda(h)=0$ for $2\eta \leqslant |h| \leqslant \pi$. If

$$\lambda(h) \sim \tfrac{1}{2}\beta_0+\sum \beta_n \cos nh, \tag{4.1}$$

then $\beta_n = O(n^{-3})$ and $\sum n|\beta_n| < \infty$. Also $\alpha_n = o(1)$. It follows from Theorem 257 that if

$$\tfrac{1}{2}\gamma_0+\sum \gamma_n \cos nh \tag{4.2}$$

is the formal product of (3.1) and (4.1), then

$$\tfrac{1}{2}\gamma_0+\sum_1^N \gamma_n \cos nh-\lambda(h)\Big\{\tfrac{1}{2}\alpha_0+\sum_1^N \alpha_n \cos nh\Big\} \to 0 \tag{4.3}$$

uniformly in h. Since (3.1) converges to $F(h)$ for $|h| \leqslant 2\eta < \delta$, and $\lambda(h)=0$ for $2\eta \leqslant |h| \leqslant \pi$, it follows from (4.3) that (4.2) converges for all h, and to a sum $F^*(h)$ defined by

$$F^*(h) = F(h)\ (|h|\leqslant \eta), \quad \lambda(h)F(h)\ (\eta \leqslant |h| \leqslant 2\eta), \quad 0\ (2\eta \leqslant |h| \leqslant \pi).$$

Since $F^*(h)$ is bounded, (4.2) is the Fourier series of $F^*(h)$.‡

If $\quad \gamma_n = -\tfrac{1}{2}n^{-2}c_n \quad (n>0), \qquad \gamma_0 = -2\sum \gamma_n = \sum n^{-2}c_n,$

then (4.2) is related to $\sum c_n$ as (3.1) is to $\sum a_n$. Since

$$F^*(h) = F(h) = o(h^2)$$

for small h, $\sum c_n$ is summable (R, 2) to 0. From this, and our assumption of the theorem in case (A), it follows that $\sum c_n$ is summable (A) to 0.

† The notation is different from that of § 1.

‡ See, e.g., *HR*, 89.

On the other hand, since $\lambda(h) = 1$ for $|h| < \eta$, it follows from (4.3) that

$$\tfrac{1}{2}(\gamma_0-\alpha_0)+\sum(\gamma_n-\alpha_n)\cos nh$$

converges uniformly to 0 for $|h| \leqslant \eta$, and that $\sum(c_n-a_n)$ is summable (R, 2) to 0. From this, and our assumption of the truth of the theorem in case (B), it follows that $\sum(c_n-a_n)$ is summable (A) to 0. Finally, since $a_n = c_n-(c_n-a_n)$, $\sum a_n$ is summable (A) to 0.

5. It remains to prove the theorem in the two special cases.

(A). In this case (3.1) is the Fourier series of a function $\phi(h)$. Since it converges to $F(h)$ for small h, $\phi(h) = F(h)$ for almost all such h, and we may suppose that this is true for all such h. Hence $\phi(h) = o(h^2)$. Now

$$\tfrac{1}{2}\alpha_0+\sum\alpha_n\cos nh\, r^n = \frac{1}{2\pi}\int\frac{1-r^2}{1-2r\cos(\theta-h)+r^2}\phi(\theta)\,d\theta$$

for $r < 1$, the limits of integration being $-\pi$ and π. Differentiating twice with respect to h, and then putting $h = 0$, we obtain

$$\sum a_n r^n = -2\sum n^2\alpha_n r^n = -\frac{2}{\pi}\int\frac{r(1-r^2)Q}{P^3}\phi(\theta)\,d\theta,$$

where

(5.1) $$P = 1-2r\cos\theta+r^2, \qquad Q = (1+r^2)\cos\theta-2r(1+\sin^2\theta).$$

Now $$|Q| \leqslant H_1\{(1-r)^2+\theta^2\}, \qquad P \geqslant H_2\{(1-r)^2+\theta^2\}$$

for appropriate H_1, H_2, and $\phi(\theta) = o(\theta^2)$. Hence

$$\sum a_n r^n = o\left(\int\frac{(1-r)\theta^2}{\{(1-r)^2+\theta^2\}^2}\,d\theta\right) = o\left(\int\frac{t^2\,dt}{(1+t^2)^2}\right) = o(1),$$

when $r \to 1$. This proves the theorem in case (A).

6. Passing to case (B), we start from the formulae

$$r^n = \frac{1}{2\pi}\int\frac{1-r^2}{P}\cos n\theta\,d\theta = \frac{1-r^2}{2\pi n^2}\int(1-\cos n\theta)\frac{\partial^2}{\partial\theta^2}\left(\frac{1}{P}\right)d\theta$$
$$= -\frac{r(1-r^2)}{\pi n^2}\int(1-\cos n\theta)\frac{Q}{P^3}\,d\theta,$$

where P and Q are defined as in (5.1). It follows that

$$\sum a_n r^n = -\frac{r(1-r^2)}{\pi}\sum\frac{a_n}{n^2}\int(1-\cos n\theta)\frac{Q}{P^3}\,d\theta.$$

Now, by hypothesis,

$$\tfrac{1}{2}\alpha_0+\sum\alpha_n\cos n\theta = \frac{1}{2}\sum\frac{a_n}{n^2}-\frac{1}{2}\sum\frac{a_n}{n^2}\cos n\theta = \frac{1}{2}\sum\frac{a_n}{n^2}(1-\cos n\theta)$$

converges uniformly to 0 for small θ, say for $|\theta| \leqslant \zeta$, so that the contribution of the integral over $(-\zeta, \zeta)$ is 0. It is thus sufficient to prove that

$$(6.1) \qquad \sum \frac{a_n}{n^2} \int_{\zeta}^{\pi} (1-\cos n\theta) \frac{Q}{P^3}\, d\theta = o\left(\frac{1}{1-r}\right)$$

for any fixed positive ζ (and *a fortiori* to prove it bounded). First,

$$\left(\sum \frac{a_n}{n^2}\right)\left(\int_{\zeta}^{\pi} \frac{Q}{P^3}\, d\theta\right)$$

is plainly bounded. Secondly, Q/P^3 and its derivatives of the first two orders are uniformly bounded in (ζ, π), and

$$j_n = \int_{\zeta}^{\pi} \cos n\theta \frac{Q}{P^3}\, d\theta = -\frac{\sin n\zeta}{n} \frac{Q(\zeta)}{P^3(\zeta)} + O\left(\frac{1}{n^2}\right),$$

by two partial integrations; and $a_n = o(n^2)$. Hence

$$(6.2) \qquad \sum \frac{a_n j_n}{n^2} = -\sum \frac{a_n \sin n\zeta}{n^3} \frac{Q(\zeta)}{P^3(\zeta)} + \sum O\left(\frac{1}{n^2}\right).$$

Since $\sum n^{-2}a_n \cos n\theta$ converges uniformly for $|\theta| \leqslant \zeta$, $\sum n^{-3}a_n \sin n\zeta$ is convergent. Hence (6.2), and so (6.1), is bounded, and this completes the proof of the theorem.

7. We need only say a few words about Theorems 255 and 256. In each case the first stage of the proof is like that of the proof of Theorem 254. In Theorem 255 our data concern the series $\sum n^{-1}a_n \sin nh$, and we must use the form assumed by Theorem 257 when A is a sine and B a cosine series. In this case the second part of the proof is trivial. If

$$(7.1) \qquad \sum n^{-1}a_n \sin nh = F(h) = o(h),$$

and the series is a Fourier series, then

$$(7.2) \qquad \sum \frac{a_n}{n^2}(1-\cos nh) = \int_0^h F(t)\, dt = o(h^2),$$

because a Fourier series can be integrated term by term; and if (7.1) converges uniformly to 0 for small h, then (7.2) is 0 for small h.

In Theorem 256 we are given (again taking $s = 0$) that

$$(7.3) \qquad \sum n^{-2}s_n \sin^2 \tfrac{1}{2}nh$$

converges to $F(h) = o(h)$ for small h, and have to prove that $\sum a_n r^n \to 0$, i.e. that $(1-r)\sum s_n r^n \to 0$. We write (7.3) in the form (3.1), with

$$\alpha_n = -\tfrac{1}{2}n^{-2}s_n \quad (n > 0), \qquad \alpha_0 = \sum n^{-2}s_n,$$

and argue as in the proof of Theorem 254. The argument proceeds much as in §§ 4–6, except that $\phi(\theta)$ is now $o(\theta)$ instead of $o(\theta^2)$, and that the conclusion is

$$\sum n^2\alpha_n r^n = o\left(\frac{1}{1-r}\right),$$

instead of $o(1)$.

8. We add in conclusion that Kuttner† has proved stronger theorems, viz. that $(R, 2) \to (C, 2+\delta)$ and $(R, 1) \to (C, 1+\delta)$ for every positive δ. There are also theorems bearing in the opposite direction. Thus Hardy and Littlewood‡ proved that $(C, -\delta) \to (R, 1)$, and Bosanquet, Paley, and Verblunsky§ have proved theorems which include both this and $(C, 1-\delta) \to (R, 2)$. The special cases $(C, 0) \to (R, 2)$ and $(C, -1) \to (R, 1)$ are familiar, the first (due to Riemann) being the regularity theorem for $(R, 2)$ summability, and the second a theorem of Fatou generalized by Hardy and Littlewood,‖ who proved that $\sum a_n = s$ $(R, 1)$ whenever $\sum a_n$ converges to s and $a_n > -H/n$.

† *PLMS* (2), 38 (1935), 273–83.

‡ *PLMS* (2), 28 (1928), 301–11 (305).

§ Bosanquet, *PLMS* (2), 31 (1930), 144–64; Paley, *PCPS*, 26 (1930), 173–203; Verblunsky, ibid. 34–42.

‖ *JLMS*, 1 (1926), 19–25.

APPENDIX IV

On Lambert and Ingham summability

1. We shall say that $\sum a_n$ is summable (L) to s, or that

$$s_n = a_1+a_2+...+a_n \to s \ \ (\mathrm{L})$$

if

$$F(y) = \sum a_n \frac{nye^{-ny}}{1-e^{-ny}} \to s \tag{1.1}$$

when $y \to +0$.† If the series in (1.1) converges for $y > 0$, then

$$F(y) = \sum s_n \Delta \frac{nye^{-ny}}{1-e^{-ny}} = -\sum s_n \int\limits_n^{n+1} \frac{d}{dt}\left(\frac{yte^{-yt}}{1-e^{-yt}}\right) dt$$

$$= y \int s(t)g(yt)\, dt = \frac{1}{x} \int g\left(\frac{t}{x}\right)s(t)\, dt,$$

where

$$g(t) = -\frac{d}{dt}\left(\frac{te^{-t}}{1-e^{-t}}\right), \qquad s(t) = \sum_{n \leqslant t} a_n, \qquad x = \frac{1}{y} \to \infty,$$

and we can also write (1.1) in the form

$$\frac{1}{x} \int g\left(\frac{t}{x}\right)s(t)\, dt \to s. \tag{1.2}$$

We saw in § 12.9(7) that $g(t)$ is W. Hence, as a corollary of Theorem 233, we obtain (A) *if* $\sum a_n$ *is summable* (L) *to* s, *and* s_n *is bounded and slowly oscillating, or real, bounded, and slowly decreasing, then* $\sum a_n$ *converges to* s. It is this theorem, with $a_n = n^{-1}\mu(n)$, which is used in the proof of the prime number theorem sketched in the note on § 12.11. The proposition (A) is sufficient for the application, and this is enough to show the interest of Lambert summability; but (A) is imperfect as it stands, and we do not state it as a formal theorem. Actually, as we shall see in a moment, the condition 's_n is bounded' is unnecessary, being a consequence of the other hypotheses.

2. There are two theorems of inclusion connecting the 'Lambert' method with more familiar methods.

THEOREM 259. *If* $\sum a_n = s$ (C, k), *for some* k, *then* $\sum a_n = s$ (L).

† The phrase was introduced by Ananda Rau, *PLMS* (2), 19 (1919), 1–20. It is convenient to begin the series with a_1, and $\sum$ will stand for $\sum\limits_1^\infty$ throughout this appendix.

Thus
$$(\mathrm{C}, k) \rightarrow (\mathrm{L}).$$
In particular, (L) is regular. The proof is straightforward,† and we do not write it out in detail. It depends on a 'convergence factor' theorem. *If*
$$\sum a_n = s \quad (\mathrm{C}, k);$$
$f_n(y)$ is continuous, and $f_n(0) = 1$, for each n; $n^k f_n(y) \rightarrow 0$ when $y > 0$ and $n \rightarrow \infty$; and $\sum n^k |\Delta^{k+1} f_n(y)| < H$, where H is independent of y; then $\sum a_n f_n(y) \rightarrow s$ when $y \rightarrow 0$. In particular this is true if
$$f_n(y) = f(ny), \qquad f(0) = 1, \qquad f(t) = o(t^{-k})$$
for large t and every k, and
$$\int t^k |f^{(k+1)}(t)|\, dt < \infty$$
for every k. Here we take $f(t) = te^{-t}/(1-e^{-t})$.

The second theorem is more difficult.

THEOREM 260. *If $\sum a_n = s$ (L), then $\sum a_n = s$ (A).*‡

Thus
$$(\mathrm{C}, k) \rightarrow (\mathrm{L}) \rightarrow (\mathrm{A}).$$
It is plain that Theorem 260 enables us to deduce a Tauberian theorem concerning summability (L) from any one concerning summability (A). In particular, after Theorem 106, we have

THEOREM 261. *If $\sum a_n = s$ (L), and s_n is slowly oscillating, or real and slowly decreasing, then $\sum a_n$ converges to s.*

This is (A) of § 1, relieved of the superfluous condition that s_n is bounded.

Theorem 260, though a pure Abelian theorem, lies deeper than (A), its proof demanding the assumption of the prime number theorem and indeed of rather more. Thus, though (A) is a corollary of Theorem 261, and the prime number theorem one of (A), we cannot base the proof of the prime number theorem on Theorem 261. We shall assume that, if

(2.1)
$$N(x) = \sum_{n \leq x} \frac{\mu(n)}{n},$$
then

(2.2)
$$\sum n^{-1} |N(n)| < \infty.$$

† The theorem is proved by Hardy, *PLMS* (2), 13 (1913), 192–8. His proof may be simplified by using the theorem quoted, which is due to Bromwich, *MA*, 65 (1908), 350–69 (358–9).

‡ Hardy and Littlewood, *PLMS* (2), 19 (1919), 21–9.

This is true if, as is known,

(2.3) $$N(x) = O\{(\log x)^{-2}\}.\dagger$$

We suppose $s = 0$, and write, for $y > 0$,

(2.4) $$f(y) = \sum a_n e^{-ny}, \qquad g(y) = \sum a_n \frac{nye^{-ny}}{1-e^{-ny}}.$$

Then

(2.5) $$g(y) = \sum_n nya_n \sum_m e^{-mny} = y\sum_m \sum_n na_n e^{-mny} = -y\sum_m f'(my),$$

the inversion being justified by absolute convergence. Hence

$$-f'(y) = \sum \mu(m)\frac{g(my)}{my} = \frac{1}{y}\sum \frac{\mu(m)}{m} g(my),\ddagger$$

$$f(y) = -\int_y^\infty f'(t)\,dt = \sum \frac{\mu(m)}{m}\int_y^\infty \frac{g(mt)}{t}\,dt = \sum \frac{\mu(m)}{m}\int_{my}^\infty \frac{g(u)}{u}\,du$$

$$= \sum N(m)\int_{my}^{(m+1)y} \frac{g(u)}{u}\,du = \sum_{m \leqslant y^{-\frac{1}{2}}} + \sum_{m > y^{-\frac{1}{2}}} = S_1 + S_2,$$

say. Now $g(u) = O(1)$, and $g(u) = o(1)$ for small u. Hence

$$S_1 = o\left\{\sum_{m \leqslant y^{-\frac{1}{2}}} |N(m)| \int_{my}^{(m+1)y} \frac{du}{u}\right\} = o\left\{\sum \frac{|N(m)|}{m}\right\} = o(1),$$

$$S_2 = O\left\{\sum_{m > y^{-\frac{1}{2}}} |N(m)| \int_{my}^{(m+1)y} \frac{du}{u}\right\} = O\left\{\sum_{m > y^{-\frac{1}{2}}} \frac{|N(m)|}{m}\right\} = o(1),$$

and $f(y) \to 0$.

3. Theorems 259 and 260 are Abelian theorems stating relations of pure inclusion. It is natural to ask when summability (L) can be inferred from summability (A), and theorems of this kind have a Tauberian character. The Tauberian condition will be an additional condition on $f(y)$.

† Actually $N(x) = O\{(\log x)^{-k}\}$ for every k: see Landau, *Handbuch*, 570, 593–7.

It is obvious that (2.3) asserts more than (*a*) $N(x) = o(1)$, and so (as we shall see in § 6) more than the prime number theorem, but it is a little less obvious that (2.2) carries these corollaries. It is, however, easily verified that (*a*) is a corollary, not only of (2.2), but of either of the weaker hypotheses.

(*b*) $\sum n^{-1}N(n)$ is convergent, (*c*) $\sum_{n \leqslant x} N(n) = o(x)$,

the second of which follows from the first by Theorem 26. In fact (*c*) is $N(n) \to 0\,(C, 1)$, and $N(n)$, being the partial sum of a series whose terms are $O(n^{-1})$, is slowly oscillating. Hence, by Theorem 68, $N(n) \to 0$.

‡ By one of the inversion formulae of Möbius: see, for example, Hardy and Wright, 237.

THEOREM 262.† *If* $\sum a_n = s$ (A), $f(y)$ *is defined by* (2.4), *and*

$$f''(y) > -Hy^{-2}, \tag{3.1}$$

then $\sum a_n = s$ (L).

We suppose $s = 0$. We have

$$f(y) = -\int_y^\infty f'(t)\,dt, \qquad g(y) = -y\sum_{m=1}^\infty f'(my),$$

by (2.5). Hence

$$S(y) = f(y)-g(y) = \sum \int_{my}^{(m+1)y} \{f'(my)-f'(t)\}\,dt = \sum u_m(y), \tag{3.2}$$

say. Also

$$yf'(y) = y\sum (f'(my)-f'\{(m+1)y\}) = \sum \int_{my}^{(m+1)y} (f'(my)-f'\{(m+1)y\})\,dt,$$

(3.3)

$$T(y) = f(y)-yf'(y)-g(y) = \sum \int_{my}^{(m+1)y} (f'\{(m+1)y\}-f'(t))\,dt = \sum v_m(y),$$

say. Now $f(y) \to 0$ and $f''(y) > -H/y^2$, and so $yf'(y) \to 0$, by Theorem 101. Hence, in order to prove that $g(y) \to 0$, it is sufficient to prove that

(3.4) $\overline{\lim}\, S(y) \leqslant 0$, (3.5) $\underline{\lim}\, T(y) \geqslant 0$.

We write

$$S(y) = \Big(\sum_{m=1}^{M} + \sum_{m=M+1}^{\infty}\Big) u_m(y) = S_1(y)+S_2(y), \tag{3.6}$$

and choose M so that $H/M < \epsilon$. If $my \leqslant t \leqslant (m+1)y$ then

$$f'(my)-f'(t) = -(t-my)f''(\tau),$$

where τ lies between the same limits. It follows that

$$f'(my)-f'(t) \leqslant Hm^{-2}y^{-1},$$

and so that

$$S_2(y) \leqslant \sum_{m>M} Hm^{-2} < HM^{-1} < \epsilon. \tag{3.7}$$

Also $yf'(y) \to 0$ with y, and $u_m(y) = o(y.y^{-1}) = o(1)$ for each m. Thus $S_1(y) \to 0$ when M is fixed; and this and (3.7) give

$$\overline{\lim}\, S(y) \leqslant \epsilon.$$

Since ϵ is arbitrary, this is (3.4).

† Hardy and Littlewood, *PLMS* (2), 41 (1936), 257–70 (258–60).

To prove (3.5), we argue similarly with $T(y)$ and $v_m(y)$. In this case, when $my \leqslant t \leqslant (m+1)y$, we have

$$f'\{(m+1)y\}-f'(t) = \{(m+1)y-t\}f''(\tau) \geqslant -Hm^{-2}y^{-1},$$

and the rest of the proof follows the same course.

If $a_n > -H/n$, then

$$f''(y) = \sum n^2 a_n e^{-ny} > -H \sum ne^{-ny} > -Hy^{-2},$$

so that the condition of Theorem 262 is certainly satisfied.

A second theorem in this direction is

THEOREM 263.† *If* $\sum a_n = s$ (A) *and* $|f'(y)| < \phi(y)$, *where* $\phi(y)$ *is a positive and decreasing function integrable down to* 0, *then* $\sum a_n = s$ (L).

Ingham's method

4. Ingham‡ has defined a method of summation which is related to his proof of the prime number theorem (§ 12.11) much as the Lambert method is related to Wiener's. We shall say that $\sum a_n$ is summable (I) to s if

$$t(x) = \sum_{1 \leqslant n \leqslant x} \frac{n}{x}\left[\frac{x}{n}\right] a_n \to s \tag{4.1}$$

when $x \to \infty$. The method is not regular, but its relations to the Cesàro methods are interesting. In particular Ingham has proved that

$$(\mathrm{C}, -\delta) \to (\mathrm{I}) \to (\mathrm{C}, \delta)$$

for every positive δ. Here we prove only

THEOREM 264. *If* $\sum a_n = s$ (I), *then* $\sum a_n = s$ (C, 1).§

THEOREM 265. *If* $\sum a_n = s$ (I), *then* $a_n = o(\log\log n)$.

We take $s = 0$ and write

$$b_n = na_n, \qquad B(x) = \sum_{n \leqslant x} b_n, \qquad F(x) = xt(x) = o(x). \tag{4.2}$$

Then we have

$$F(x) = \sum_{n \leqslant x} b_n \sum_{m \leqslant x/n} 1 = \sum_{m \leqslant x} \sum_{n \leqslant x/m} b_n = \sum_{m \leqslant x} B\left(\frac{x}{m}\right), \tag{4.3}$$

† Ananda Rau, l.c. (Theorem 22). His condition (ii) is superfluous, being a corollary of (i). The short proof given by Hardy and Littlewood, l.c. 259, is fallacious.

‡ L.c. under § 12.9. (See Corrigenda, p. 386.)

§ We shall see in § 6 that (C, -1) $\to$ (I) is a corollary of 'Axer's theorem'.

and so, by another of the Möbius formulae†

$$B(x) = \sum_{n \leqslant x} \mu(n) F\left(\frac{x}{n}\right).$$

Now

$$(4.4)\quad A^{(1)}(x) = \sum_{n \leqslant x} (x-n)a_n = \sum_{n \leqslant x} \left(\frac{x}{n}-1\right)b_n = \int_0^x \left(\frac{x}{u}-1\right) dB(u)$$

$$= -\int_0^x B(u)\, d\left(\frac{x}{u}-1\right) = x\int_1^x \frac{B(u)}{u^2}\, du,$$

$$\int_1^x \frac{B(u)}{u^2}\, du = \int_1^x \left\{\sum_{n \leqslant u} \mu(n) F\left(\frac{u}{n}\right)\right\} \frac{du}{u^2} = \sum_{n \leqslant x} \mu(n) \int_n^x F\left(\frac{u}{n}\right) \frac{du}{u^2}$$

$$= \sum_{n \leqslant x} \frac{\mu(n)}{n} \int_1^{x/n} F(v) \frac{dv}{v^2} = \int_1^x \frac{F(v)}{v^2} \left\{\sum_{n \leqslant x/v} \frac{\mu(n)}{n}\right\} dv = \int_1^x \frac{F(v)}{v^2} N\left(\frac{x}{v}\right) dv,$$

where $N(x)$ is defined as in § 2: we make the same assumption about $N(x)$ as there. It follows that

$$\int_1^x \frac{B(u)}{u^2}\, du = \frac{1}{x}\int_1^x F\left(\frac{x}{w}\right) N(w)\, dw = \frac{1}{x}\left(\int_1^{\sqrt{x}} + \int_{\sqrt{x}}^x\right).$$

After (4.2), this is

$$\frac{1}{x}\int_1^{\sqrt{x}} o\left(\frac{x}{w}\right) |N(w)|\, dw + \frac{1}{x}\int_{\sqrt{x}}^x O\left(\frac{x}{w}\right) |N(w)|\, dw$$

$$= o\left\{\int_1^{\sqrt{x}} \frac{|N(w)|}{w}\, dw\right\} + O\left\{\int_{\sqrt{x}}^x \frac{|N(w)|}{w}\, dw\right\} = o(1).$$

Thus $A^{(1)}(x) = o(x)$, by (4.4), i.e. $\sum a_n = 0$ (C, 1).

This proves Theorem 264, and we can now transfer Tauberian theorems for summability (C) to summability (I). In particular we have

THEOREM 266. *If $\sum a_n = s$ (I), and s_n is slowly oscillating, or real and slowly decreasing, then $\sum a_n$ converges to s.*

To prove Theorem 265, we write $F(x)$ in the form

$$F(x) = \sum_{mn \leqslant x} b_n = \sum_{q \leqslant x} \sum_{n|q} b_n = \sum_{q \leqslant x} \beta_q,$$

† Hardy and Wright, 236.

say. Since $F(x) = o(x)$, $\beta_q = o(q)$. Hence, using the most familiar of the Möbius formulae,†

$$qa_q = b_q = \sum_{d|q} \mu\left(\frac{q}{d}\right)\beta_d = o\left(\sum_{d|q} d\right) = o\{\sigma(q)\},$$

where $\sigma(q)$ is the sum of the divisors of q. But $\sigma(q) = O(q \log\log q)$,‡ and so $a_q = o(\log\log q)$.

The deduction of the prime number theorem from (12.11.4)

5. We have taken it for granted, in § 12.11 and elsewhere, that the prime number theorem can be deduced from (12.11.4) by elementary reasoning. This was proved by Landau in 1911,§ but the proof is not given in any book. The deduction depends on an important elementary theorem which we shall call 'Axer's theorem'.‖

THEOREM 267. *If (a) $\chi(x)$ is of bounded variation in every finite interval $1 \leqslant x \leqslant X$,*

$$(b) \quad \sum_{n \leqslant x} a_n = o(x),$$

and **either** *of the pairs of conditions*

$$(c\,1) \quad \chi(x) = O(1), \qquad (d\,1) \quad \sum_{n \leqslant x} |a_n| = O(x),$$

$$(c\,2) \quad \chi(x) = O(x^\alpha) \quad (0 < \alpha < 1), \qquad (d\,2) \quad a_n = O(1),$$

is satisfied, then

$$\sum_{n \leqslant x} a_n \chi\left(\frac{x}{n}\right) = o(x).$$

We suppose $0 < \delta < 1$, and write

$$S(x) = \sum_{n \leqslant x} a_n \chi\left(\frac{x}{n}\right) = \sum_{n \leqslant \delta x} + \sum_{\delta x < n \leqslant x} = S_1 + S_2, \tag{5.1}$$

say. Then, if $A_0 = 0$, $a_1 + a_2 + \ldots + a_n = A_n$, we have

$$S_2 = \sum_{\delta x < n \leqslant x} (A_n - A_{n-1})\chi\left(\frac{x}{n}\right) = \sum_{n=[\delta x]+1}^{[x]} (A_n - A_{n-1})\chi\left(\frac{x}{n}\right)$$

$$= -A_{[\delta x]}\chi\left(\frac{x}{[\delta x]+1}\right) + A_{[x]}\chi\left(\frac{x}{[x]}\right) + \sum_{n=[\delta x]+1}^{[x]-1} A_n\left\{\chi\left(\frac{x}{n}\right) - \chi\left(\frac{x}{n+1}\right)\right\}.$$

Now

$$\sum_{[\delta x]+1}^{[x]-1} \left|\chi\left(\frac{x}{n}\right) - \chi\left(\frac{x}{n+1}\right)\right|, \qquad \left|\chi\left(\frac{x}{[\delta x]+1}\right)\right|, \qquad \left|\chi\left(\frac{x}{[x]}\right)\right|$$

† Hardy and Wright, 235. ‡ Ibid. 264.

§ *WS*, 120 (1911), 973–88. For the converse inference see Landau, *Handbuch*, 588–90.

‖ The actual theorem proved by Axer [*PMF*, 21 (1910), 65–95] is Theorem 267, with $\chi(x) = x - [x]$ and a_n subject to (*b*) and (*d* 1). It is the second form of the theorem, with conditions (*c* 2) and (*d* 2), which is most important here. The arrangement and proof are Ingham's. (But see Corrigenda, p. 386.)

are, after (a), all less than functions of δ only. Thus

$$|S_2| \leqslant o(x)P(\delta), \tag{5.2}$$

where $P(\delta)$ depends only on δ.

If conditions (c 1) and (d 1) are satisfied, then

$$|S_1| \leqslant H \sum_{n \leqslant \delta x} |a_n| \leqslant H\delta x$$

for a fixed H. If conditions (c 2) and (d 2) are satisfied, then

$$|S_1| \leqslant H \sum_{n \leqslant \delta x} \left(\frac{x}{n}\right)^\alpha \leqslant Hx^\alpha(\delta x)^{1-\alpha} = H\delta^{1-\alpha}x.$$

In either case we can choose $\delta(\epsilon)$ so that $|S_1| < \epsilon x$; and then, after (5.2), choose $x_0 = x_0(\delta, \epsilon) = x_0(\epsilon)$ so that $|S_2| < \epsilon x$ for $x \geqslant x_0$. It follows that $S(x) = o(x)$.

6. The prime number theorem is equivalent to $\psi(x) \sim x$, where

$$\psi(x) = \sum_{n \leqslant x} \Lambda(n) = \sum_{p^m \leqslant x} \log p; \tag{6.1}$$

and

$$\sum \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)}$$

for $\Re s > 1$.† Thus if

$$H(s) = -\frac{\zeta'(s)}{\zeta(s)} - \zeta(s) + 2\gamma = \sum \frac{c_n}{n^s},$$

then

$$c_1 = 2\gamma - 1, \qquad c_n = \Lambda(n) - 1 \quad (n > 1),$$

$$C(x) = \sum_{n \leqslant x} c_n = \sum_{n \leqslant x} \{\Lambda(n) - 1\} + 2\gamma = \psi(x) - [x] + 2\gamma,$$

and the prime number theorem is equivalent to $C(x) = o(x)$. Also

$$H(s) = \frac{1}{\zeta(s)}\{-\zeta'(s) - \zeta^2(s) + 2\gamma\zeta(s)\} = F(s)G(s),$$

where

$$F(s) = \frac{1}{\zeta(s)} = \sum \frac{a_n}{n^s}, \qquad a_n = \mu(n), \tag{6.2}$$

$$G(s) = -\zeta'(s) - \zeta^2(s) + 2\gamma\zeta(s) = \sum \frac{b_n}{n^s}, \tag{6.3}$$

so that $b_n = \log n - d(n) + 2\gamma$, where $d(n)$ is the number of divisors of n.‡ Thus, assuming (12.11.4), we have

$$A(x) = \sum_{n \leqslant x} a_n = \sum_{n \leqslant x} \mu(n) = M(x) = o(x), \tag{6.4}$$

$$B(x) = \sum_{n \leqslant x} b_n = \sum_{n \leqslant x} \log n - \sum_{n \leqslant x} d(n) + 2\gamma \sum_{n \leqslant x} 1.$$

† See Hardy and Wright, 252, 344 et seq.

‡ Ibid., 249.

The first sum in $B(x)$ is $x\log x-x+O(\log x)$; the second is

$$x\log x+(2\gamma-1)x+O(\sqrt{x}),$$

by a familiar result in the theory of 'Dirichlet's divisor problem';† and the third is $x+O(1)$. It follows that

(6.5) $$B(x)=O(\sqrt{x}).$$

Now from $H(s)=F(s)G(s)$, i.e. $\sum c_n n^{-s}=\sum a_n n^{-s}\sum b_n n^{-s}$, it follows that

$$C(x)=\sum_{n\leqslant x}c_n=\sum_{mn\leqslant x}a_m b_n=\sum_{m\leqslant x}a_m\sum_{n\leqslant x/m}b_n=\sum_{m\leqslant x}a_m B\left(\frac{x}{m}\right).$$

Here (1) $B(x)$ is of bounded variation in any finite interval $(1,X)$; (2) $A(x)=o(x)$, by (6.4); (3) $B(x)=O(\sqrt{x})$, by (6.5); and (4) $a_m=O(1)$. Thus the conditions (a), (b), (c 2), and (d 2) of Theorem 267 are satisfied, and so $C(x)=o(x)$. Thus the prime number theorem is a corollary of (6.4), and *a fortiori* of the convergence of $\sum n^{-1}\mu(n)$.‡

Here we have used the second form of Theorem 267. As an example of the use of Axer's theorem in its original form, we prove the implication $(C,-1)\to(I)$. Rather more generally, we prove

THEOREM 268. *If $\sum a_n$ converges to s, and $na_n>-H$, then $\sum a_n$ is summable* (I) *to s.*

For then (taking $s=0$)

(6.6) $$\sum_{n\leqslant x}na_n=o(x),\qquad \sum_{n\leqslant x}n|a_n|=\sum_{n\leqslant x}na_n-2\sum_{n\leqslant x}na_n^-=O(x),$$

and therefore, by Theorem 267,

(6.7) $$\sum_{n\leqslant x}na_n\left(\frac{x}{n}-\left[\frac{x}{n}\right]\right)=o(x),$$

$$\sum_{n\leqslant x}\frac{n}{x}\left[\frac{x}{n}\right]a_n=\sum_{n\leqslant x}a_n+o(1)=o(1).$$

As a further example, if $a_n=n^{-1}\mu(n)$, then the first of (6.6) is (6.4) and the second is trivial, and (6.7) follows, by Theorem 267. Also

$$\sum_{n\leqslant x}\mu(n)\left[\frac{x}{n}\right]=\sum_{mn\leqslant x}\mu(n)=\sum_{q\leqslant x}\sum_{n|q}\mu(n)=1,$$

since the inner sum is 1 when $q=1$ and 0 otherwise. It follows from (6.7) that

$$x\sum_{n\leqslant x}\frac{\mu(n)}{n}=o(x),$$

i.e. that $\sum n^{-1}\mu(n)$ converges to 0. Thus this is an elementary consequence of (6.4).

† Hardy and Wright, 262.

‡ See the footnote to p. 374. On the other hand, the prime number theorem implies $\Sigma n^{-1}\mu(n)=0$, Landau, *Handbuch*, 591–3.

APPENDIX V

Two theorems of M. L. Cartwright†

1. If $\sum a_n = s\ (\mathrm{A}, p)$, i.e. if

(1.1) $$\sum a_n e^{-yn^p} \to s$$

when $y \to 0$, and $0 < q < p$, then $\sum a_n = s\ (\mathrm{A}, q)$, i.e.

(1.2) $$\sum a_n e^{-yn^q} \to s,$$

provided only that the series (1.2) converges for $y > 0$. We prove this here, in a more complete form and with a companion theorem in the opposite direction.

If $y = re^{i\theta}$ and (1.1) holds uniformly in the angle

(1.3) $$-\tfrac{1}{2}\pi < \alpha_1 \leqslant \theta \leqslant \alpha_2 < \tfrac{1}{2}\pi,$$

then we shall say that

(1.4) $$\sum a_n = s\ \ (\mathrm{A}, p, \alpha_1, \alpha_2).$$

We shall prove

THEOREM 269. *If*

(1.5) $$p > 0, \quad q = kp, \quad 0 < k < 1,$$

(1.4) *is true, and* $\sum a_n e^{-yn^q}$ *is convergent for* $y > 0$, *then*

(1.6) $$\sum a_n = s\ \ (\mathrm{A}, q, \beta_1, \beta_2)$$

for

(1.7) $$-\tfrac{1}{2}\pi + k(\tfrac{1}{2}\pi + \alpha_1) < \beta_1 < \beta_2 < \tfrac{1}{2}\pi - k(\tfrac{1}{2}\pi - \alpha_2).$$

THEOREM 270. *If*

(1.8) $$p > 0, \quad q = kp, \quad k > 1,$$

(1.9) $$\alpha_2 - \alpha_1 > \pi\left(1 - \frac{1}{k}\right),$$

and (1.4) *is true, then* (1.6) *is true under the conditions* (1.7).

Some preliminary remarks are desirable.

(1) In Theorem 269, $0 < q < p$, and the theorem includes that cited at the beginning of the section. The convergence of $\sum a_n e^{-yn^q}$, for $y > 0$, must be taken as a hypothesis. In Theorem 270, $q > p$. The convergence of $\sum a_n e^{-yn^p}$ for $y > 0$, implied in (1.1), carries with it that of $\sum a_n e^{-yn^q}$, so that no hypothesis concerning this is needed.

† M. L. Cartwright, *PLMS* (2), 31 (1930), 81–96 (where the method of proof is different).

(2) It is plain that the range (β_1, β_2) prescribed by (1.7) always lies inside $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$. If $k < 1$ then

$$-\tfrac{1}{2}\pi + k(\tfrac{1}{2}\pi + \alpha_1) - \alpha_1 = -(1-k)(\tfrac{1}{2}\pi + \alpha_1) < 0,$$

so that $-\frac{1}{2}\pi + k(\frac{1}{2}\pi + \alpha_1) < \alpha_1$; and similarly $\frac{1}{2}\pi - k(\frac{1}{2}\pi - \alpha_2) > \alpha_2$. Thus (β_1, β_2) may extend beyond (α_1, α_2) at each end: the angle of summability in the conclusion is greater than that in the hypothesis. We may suppose $\alpha_1 = \alpha_2$, in which case the hypothesis asserts (1.1) along one line only.

If $k > 1$ then $-\frac{1}{2}\pi + k(\frac{1}{2}\pi + \alpha_1) > \alpha_1$ and $\frac{1}{2}\pi - k(\frac{1}{2}\pi - \alpha_2) < \alpha_2$, so that (β_1, β_2) lies inside (α_1, α_2) at both ends, and the angle of summability in the conclusion is smaller than that in the hypothesis. The hypothesis (1.9) is equivalent to

$$-\tfrac{1}{2}\pi + k(\tfrac{1}{2}\pi + \alpha_1) < \tfrac{1}{2}\pi - k(\tfrac{1}{2}\pi - \alpha_2),$$

and is essential to the conclusion. All this is in harmony with the general principle stated in § 4.12. If $q < p$ then the (A, q) method is less powerful than the (A, p) method, and therefore less likely to be 'applicable', but, if applicable, it is more effective.

2. We need a lemma concerning the Fourier transform of e^{-x^k}.

THEOREM 271. *If*

$$k > 0, \qquad \frac{\pi}{2k} = \lambda, \qquad x = re^{i\theta}, \qquad F(x) = \int_0^\infty e^{-t^k} \cos xt \, dt$$

for positive x, then (1) *the analytic function $F(x)$ defined by the integral is regular for $|\theta| < \lambda$,* (2) *$F(x) = O(1)$ for small x, and* (3) *$F(x) = O(r^{-1-k})$ for large x, each of* (2) *and* (3) *holding uniformly in any angle $|\theta| \leqslant \lambda - \epsilon < \lambda$.*

If $k > 1$, then $F(x)$ is an integral function and (1) and (2) are trivial. If $k = 1$, then $F(x) = (1+x^2)^{-1}$, the integral converging for $|\Im(x)| < 1$. If $k < 1$, then the integral for $F(x)$ is convergent only for real x. But in any case we have

$$F(x) = \frac{1}{x} \int e^{-(t/x)^k} \cos t \, dt,$$

first for positive x and then for $\cos k\theta > 0$, i.e. for $|\theta| < \lambda$. Also

$$|F(x)| \leqslant \frac{1}{r} \int e^{-(t/r)^k \cos k(\lambda - \epsilon)} \, dt$$

for $|\theta| \leqslant \lambda - \epsilon$, so that (2) is true for $k > 0$.

Next, for positive x,

$$xF(x) = x\int e^{-t^k}\cos xt\,dt = k\int e^{-t^k}t^{k-1}\sin xt\,dt$$

$$= \frac{k}{2i}\int e^{-t^k}t^{k-1}e^{ixt}\,dt - \frac{k}{2i}\int e^{-t^k}t^{k-1}e^{-ixt}\,dt = xF_1(x) - xF_2(x),$$

say. Still supposing that x is positive, we have

$$xF_1(x) = \frac{k}{2ix^k}\int e^{-(t/x)^k}t^{k-1}e^{it}\,dt;$$

and we may take the integral along the line $t = \rho e^{i\phi}$, where ϕ is small and positive. The integral is then absolutely and uniformly convergent for x in any angle $|\theta| \leqslant \lambda-\phi-\eta < \lambda-\phi$, and represents $xF_1(x)$ throughout $|\theta| < \lambda-\phi$; and

$$x^{k+1}F_1(x) \to \frac{k}{2i}\int\limits_0^{\infty e^{i\phi}} t^{k-1}e^{it}\,dt = \frac{\Gamma(1+k)}{2i}e^{\frac{1}{2}k\pi i}$$

when x tends to infinity, uniformly in $|\theta| \leqslant \lambda-\phi-\eta$. Similarly, using a path of integration below the real axis, we see that

$$2ix^{k+1}F_2(x) \to \Gamma(1+k)e^{-\frac{1}{2}k\pi i},$$

uniformly in the same angle, and (3) follows.

3. Passing to the proof of Theorems 269 and 270, we have

$$e^{-yn^q} = \frac{2}{\pi}\int F(t)\cos(y^{1/k}n^pt)\,dt = \frac{1}{\pi}\int F(t)e^{iYt}\,dt + \frac{1}{\pi}\int F(t)e^{-iYt}\,dt, \tag{3.1}$$

if $y > 0$ and $Y = y^{1/k}n^p$. If $t = \rho e^{i\phi}$, $0 \leqslant \phi_1 \leqslant \pi$ and $-\pi \leqslant \phi_2 \leqslant 0$, then $\arg(-iYt)$ and $\arg(iYt)$ lie between $-\frac{1}{2}\pi$ and $\frac{1}{2}\pi$ for $0 \leqslant \phi \leqslant \phi_1$ and $\phi_2 \leqslant \phi \leqslant 0$ respectively. If also

$$|\phi_1| < \lambda-\epsilon, \qquad |\phi_2| < \lambda-\epsilon, \tag{3.2}$$

then it follows from Theorem 271 and Cauchy's theorem that we may replace (3.1) by

$$e^{-yn^q} = \frac{1}{\pi}\int\limits_{C_1} F(t)e^{iYt}\,dt + \frac{1}{\pi}\int\limits_{C_2} F(t)e^{-iYt}\,dt, \tag{3.3}$$

where C_1 and C_2 are the radii $\phi = \phi_1$ and $\phi = \phi_2$. It also follows that (3.3), proved at present only for $y > 0$, is true for complex $y = re^{i\theta}$, and ϕ_1, ϕ_2 satisfying (3.2), and

$$-\tfrac{1}{2}\pi \leqslant -\tfrac{1}{2}\pi+\frac{\theta}{k}+\phi_1 \leqslant \tfrac{1}{2}\pi, \qquad -\tfrac{1}{2}\pi \leqslant \tfrac{1}{2}\pi+\frac{\theta}{k}+\phi_2 \leqslant \tfrac{1}{2}\pi \tag{3.4}$$

(instead of $0 \leqslant \phi_1 \leqslant \pi$, $-\pi \leqslant \phi_2 \leqslant 0$). *A fortiori* it is true if

(3.5) $$\alpha_1 \leqslant -\tfrac{1}{2}\pi+\frac{\theta}{k}+\phi_1 \leqslant \alpha_2, \qquad \alpha_1 \leqslant \tfrac{1}{2}\pi+\frac{\theta}{k}+\phi_2 \leqslant \alpha_2,$$

and ϕ_1 and ϕ_2 satisfy (3.2).

Let us assume provisionally that, when $\beta_1 \leqslant \theta \leqslant \beta_2$, we can choose $\phi_1 = \phi_1(\theta)$ and $\phi_2 = \phi_2(\theta)$ so as to satisfy (3.2) and (3.5). Let $\tau = \exp\{\tfrac{1}{2}(\alpha_1+\alpha_2)i\}$, and take $\delta > 0$. Then

(3.6)
$$\sum a_n e^{-yn^q-\delta\tau n^p} = \frac{1}{\pi}\int\limits_{C_1} \sum a_n e^{-n^p z_1} F(t)\,dt + \frac{1}{\pi}\int\limits_{C_2} \sum a_n e^{-n^p z_2} F(t)\,dt = \chi_1+\chi_2,$$

say, where $\qquad z_1 = \delta\tau - iy^{1/k}t, \qquad z_2 = \delta\tau + iy^{1/k}t,$

provided that we can justify the summations under the integral sign.

If $z_1 = \delta\tau + Z_1$, then $\arg Z_1$ lies in (α_1, α_2), and z_1 lies in the angle $D(\delta)$ whose vertex is at $\delta\tau$ and whose sides make angles α_1 and α_2 with the positive real axis. Thus, if δ, y are fixed, $\sum a_n e^{-n^p z_1}$ converges uniformly on $\arg t = \phi_1$; and

$$\int\limits_{C_1} |F(t)|\,dt$$

is convergent. It follows that the term-by-term integration in χ_1 is legitimate, and that in χ_2 may be justified similarly.

Next, $f_p(z) = \sum a_n e^{-n^p z}$ is, by hypothesis, uniformly continuous in $D(0)$, and therefore, by Theorem 31,

$$f_q(y) = \sum a_n e^{-yn^q} = \lim_{\delta\to 0} \sum a_n e^{-yn^q-\delta\tau n^p}$$
$$= \frac{1}{\pi}\int\limits_{C_1} f_p(-iy^{1/k}t)F(t)\,dt + \frac{1}{\pi}\int\limits_{C_2} f_p(iy^{1/k}t)F(t)\,dt,$$

if the series $\sum a_n e^{-yn^q}$ is convergent (a hypothesis needed only when $q < p$, $k < 1$, as we pointed out in § 1).

Finally, $f_p(-iy^{1/k}t)$ and $f_p(iy^{1/k}t)$ tend uniformly to s, when $y \to 0$ in the angle $\beta_1 \leqslant \theta \leqslant \beta_2$ and $|t|$ is bounded, and are uniformly bounded in the angle, for all $|t|$; and

$$\int\limits_{C_1} |F(t)|\,dt, \qquad \int\limits_{C_2} |F(t)|\,dt$$

are bounded and uniformly convergent for ϕ_1, ϕ_2 satisfying (3.2).

Hence $$f_q(y) \to \frac{s}{\pi}\left(\int\limits_{C_1} + \int\limits_{C_2}\right) F(t)\,dt = \frac{2s}{\pi}\int\limits_0^\infty F(t)\,dt = s$$

uniformly in $\beta_1 \leqslant \theta \leqslant \beta_2$.

It remains to justify our provisional assumption that ϕ_1 and ϕ_2 can be chosen so as to satisfy (3.2) and (3.5), and for this purpose we may take $\epsilon = 0$. Considering the inequalities for ϕ_1, we have to show that $\phi_1 = \phi_1(\theta)$ can be chosen so that

$$-\frac{\pi}{2k} < \phi_1 < \frac{\pi}{2k}, \qquad \alpha_1+\tfrac{1}{2}\pi-\frac{\theta}{k} \leqslant \phi_1 \leqslant \alpha_2+\tfrac{1}{2}\pi-\frac{\theta}{k}.$$

This is certainly possible if

$$\alpha_1+\tfrac{1}{2}\pi-\frac{\theta}{k} < \frac{\pi}{2k}, \qquad -\frac{\pi}{2k} < \alpha_2+\tfrac{1}{2}\pi-\frac{\theta}{k},$$

i.e.
$$-\tfrac{1}{2}\pi+k(\tfrac{1}{2}\pi+\alpha_1) < \theta < \tfrac{1}{2}\pi+k(\tfrac{1}{2}\pi+\alpha_2),$$

and *a fortiori* if $\beta_1 \leqslant \theta \leqslant \beta_2$. Thus ϕ_1, and similarly ϕ_2, can be chosen.

4. It may be useful to add a short statement of the principal properties of $F(x)$: Theorem 271 contains only the minimum necessary for the proof of Theorems 269 and 270. If $k > 1$ then

(4.1)
$$F(x) = \frac{1}{k}\sum \frac{(-1)^m}{(2m)!}\Gamma\left(\frac{2m+1}{k}\right)x^{2m}$$

is an integral function. It has an asymptotic expansion

(4.2)
$$\sum \frac{(-1)^{m-1}}{m!}\Gamma(mk+1)\sin\tfrac{1}{2}mk\pi\, x^{-mk-1}$$

valid in the sectors $|\arg x| < \lambda$ and $|\arg(-x)| < \lambda$. This series vanishes identically when $k = 2, 4, 6,\ldots$, and in this case $F(x)$ is exponentially small for large x in the sectors, and all its zeros are real. In any case it is exponentially large in the remaining sectors. The exponential approximations may be found by the saddle-point method.

When $k < 1$ the integral for $F(x)$ converges only for real x; but the series (4.2) is convergent, and $F(x) = x^{-1}G(x^{-k})$, where $G(w)$ is integral. In this case the series (4.1), now divergent, represents $F(x)$ asymptotically for small x with $|\arg x| < \lambda$.

When $0 < k \leqslant 2$, $F(x) > 0$ for positive x.

Fuller information, and references, will be found in Pólya, *MM*, 52 (1923), 185–8; Wright, *JLMS*, 10 (1935), 286–93; and *PTRS*, 238 (1940), 423–51, and 239 (1946), 217–32.

5. Finally, we add a few words concerning the application of the (A, q) method to Fourier series. The main theorem is

THEOREM 272. *The* (A, q) *method is Fourier-effective for every* q.

We need an extension of the ordinary theorem concerning A or (A, 1) summability, viz.

THEOREM 273. *If $f(t)$ satisfies l_c for $t = \theta$,† and $-\frac{1}{2}\pi < \alpha_1 < \alpha_2 < \frac{1}{2}\pi$, then the Fourier series of $f(t)$ is summable* $(A, 1, \alpha_1, \alpha_2)$, *for $t = \theta$, to c. In particular this is true, with $c = f(\theta)$, for almost all θ.*

We leave the proof, which is similar to the ordinary proof of A summability, to the reader.‡ Theorem 272 is a corollary. We can choose α so that

$$0 < \tfrac{1}{2}\pi - q(\tfrac{1}{2}\pi - \alpha) < \tfrac{1}{2}\pi$$

(as is always possible by taking α near enough to $\frac{1}{2}\pi$). Then the Fourier series is summable $(A, 1, -\alpha, \alpha)$, by Theorem 273; and therefore, by Theorem 269, summable $(A, q, -\beta, \beta)$ provided that

$$0 < \beta < \tfrac{1}{2}\pi - q(\tfrac{1}{2}\pi - \alpha).$$

In particular, it is summable (A, q).

† i.e. (1.2) of Appendix II.

‡ It is only necessary to show that if

$$K(y, t) = \tfrac{1}{2} + \sum e^{-ny}\cos nt = \frac{1-e^{-2y}}{2(1-2e^{-y}\cos t + e^{-2y})},$$

$$y = u + iv, \qquad v = u \tan\chi, \qquad |\chi| \leqslant \chi_0 < \tfrac{1}{2}\pi,$$

then
$$\int_{-\pi}^{\pi} \left| t\frac{\partial K}{\partial t}\right| dt < H,$$

where H depends only on χ_0. See Appendix II.

CORRIGENDA

p. 286. The statement that 'any h of L can be expressed in the form (12.2.1), with an r of L' is false, since $H(t)/G(t) = R(t) \to 0$ as $|t| \to \infty$ whenever r is L (Titchmarsh, *Fourier transforms*, Theorem 1). But in the relevant theorem (Theorem 229) h is restricted so that $H(t) = 0$ for large $|t|$.

p. 376. A summation method equivalent to Ingham's method, (I), was given earlier by A. Wintner, *Eratosthenian averages*, Baltimore, 1943.

p. 378. The extension of 'Axer's theorem' given as Theorem 267, and indeed with the more general conditions (c 2), (d 1) in place of the alternatives (c 1), (d 1) or (c 2), (d 2), is included in a theorem of E. Landau, *RP*, 34 (1912), 121–31, Satz 5.

LIST OF BOOKS

[Books not in this list are referred to under their full titles]

Abbreviated title	*Full title*	*First reference*
Andersen, *Studier*.	A. F. Andersen, *Studier over Cesàro's Summabilitets-metode* (Dissertation, Copenhagen, 1921).	p. 118, § 5.5
Bailey.	W. N. Bailey, *Generalized hypergeometric series*, Cambridge, 1935 (Cambridge tracts in mathematics, No. 32).	p. 118, § 5.8
Bohr, *Bidrag*.	H. Bohr, *Bidrag til de Dirichlet'ske Rækkers Theori* (Dissertation, Copenhagen, 1910).	p. 118, § 5.5
Borel.	E. Borel, *Leçons sur les séries divergentes*, ed. 2, Paris, 1928.	p. 40, § 2.5
Bromwich.	T. J. I'A. Bromwich, *An introduction to the theory of infinite series*, ed. 2, Cambridge, 1926. (Occasional references to the first edition of 1906 are given as Bromwich (1).)	p. 21, § 1.1
Dienes.	P. Dienes, *The Taylor Series*, Oxford, 1931.	p. 60, § 3.2
Ford, *Studies*.	W. B. Ford, *Studies on divergent series and summability* (University of Michigan Science Series, New York, 1916).	p. 40, § 2.5
Hardy.	G. H. Hardy, *A course of pure mathematics*, ed. 9, Cambridge, 1944.	p. 21, § 1.1
Hardy and Riesz.	G. H. Hardy and M. Riesz, *The general theory of Dirichlet's series*, Cambridge, 1915 (Cambridge tracts in mathematics, No. 18).	p. 92, § 4.9
Hardy and Rogosinski, *HR*.	G. H. Hardy and W. W. Rogosinski, *Fourier series*, Cambridge, 1944 (Cambridge tracts in mathematics, No. 38).	p. 40, § 2.7
Hardy and Wright.	G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, ed. 2, Oxford, 1945.	p. 92, § 4.10
Hobson, 1 & 2.	E. W. Hobson, *The theory of functions of a real variable and the theory of Fourier series*, vol. 1, ed. 3, Cambridge, 1927, and vol. 2, ed. 2, 1926.	p. 21, § 1.1
Inequalities.	G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge, 1934.	p. 282, § 11.17
Ingham.	A. E. Ingham, *The distribution of prime numbers*, Cambridge, 1932 (Cambridge tracts in mathematics, No. 30).	p. 39, § 2.2
Knopp.	K. Knopp, *Theory and application of infinite series*, English ed., London, 1928.	p. 40, § 2.5

Kogbetliantz.	E. Kogbetliantz, *Sommation des séries et intégrales divergentes par les moyennes arithmétiques et typiques*, Paris, 1931 (Mémorial des sciences mathématiques, No. 51).	p. 118, § 5.5
Koksma.	J. F. Koksma, *Diophantische Approximation*, Berlin, 1936 (Ergebnisse d. Math. 4, Heft 4).	p. 119, § 5.17
Landau, *Ergebnisse*.	E. Landau, *Darstellung und Begründung einiger neuerer Ergebnisse der Funktionentheorie*, ed. 2, Berlin, 1929.	p. 118, § 5.9
Landau, *Handbuch*.	E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, Leipzig, 1909 (two volumes paged consecutively).	p. 39, § 2.2
Lindelöf.	E. Lindelöf, *Le calcul des résidus et ses applications à la théorie des fonctions*, Paris, 1905.	p. 347, § 13.1
Littlewood.	J. E. Littlewood, *Lectures on the theory of functions*, Oxford, 1943.	p. 225, § 9.13
Moore, *Convergence factors*.	C. N. Moore, *Summable series and convergence factors*, New York, 1938 (Amer. Math. Soc. Colloquium Publications, No. 22).	p. 62, § 3.5
Pólya and Szegö.	G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, 2 vols., Berlin, 1925.	p. 41, § 2.13
Tannery and Molk.	J. Tannery and J. Molk, *Éléments de la théorie des fonctions elliptiques*, 4 vols., Paris, 1893–1902.	p. 92, § 4.10
Titchmarsh, *Fourier integrals*.	E. C. Titchmarsh, *Introduction to the theory of Fourier integrals*, Oxford, 1937.	p. 281, § 11.12
Titchmarsh, *Theory of functions*.	E. C. Titchmarsh, *The theory of functions*, ed. 2, Oxford, 1939.	p. 159, f.n.
Watson.	G. N. Watson, *A treatise on the theory of Bessel functions*, Cambridge, 1922.	p. 355, f.n.
Whittaker and Watson.	E. T. Whittaker and G. N. Watson, *A course of modern analysis*, ed. 4, Cambridge, 1927.	p. 40, § 2.2
Widder.	D. V. Widder, *The Laplace transform*, Princeton, 1941.	p. 92, § 4.13
Wiener, *The Fourier integral*.	N. Wiener, *The Fourier integral and certain of its applications*, Cambridge, 1933.	p. 316, § 12.1
Zygmund.	A. Zygmund, *Trigonometrical series*, Warsaw, 1935 (Monografje Matematyczne, No. 5).	p. 40, § 2.7

LIST OF PERIODICALS

[Periodicals not in this list are referred to under their full titles]

Abbreviated titles	*Full title*	*First reference*
AEN	Annales scientifiques de l'École normale supérieure	p. 39, § 2.2
AM	Acta Mathematica	p. 40, § 2.5
Annals	Annals of Mathematics	p. 40, § 2.8
ASF	Acta Societatis sci. Fennicae	p. 199, § 8.11
AT	Annales de la Faculté des sciences de Toulouse	p. 197, § 8.10
AUH	Acta litt. ac sci. Univ. Hungaricae (Szeged)	p. 146, § 6.4
BAMS	Bulletin of the American Mathematical Society	p. 61, § 3.4
BAP	Bulletin international de l'Académie polonaise (Cracovie)	p. 146, §§ 6.5–6
BS	Sitzungsberichte der K. Preussischen Akademie der Wissenschaften (Berlin)	p. 224, § 9.5
BSM	Bulletin des sciences mathématiques	p. 118, §§ 5.2–3
CR	Comptes rendus de l'Académie des sciences (Paris)	p. 91, §§ 4.7–8
DMJ	Duke Mathematical Journal	p. 317, § 12.12
DMV	Jahresbericht der Deutschen Math.-Vereinigung	p. 245, § 10.4
J de M	Journal de mathématiques	p. 197, § 8.10
JIMS	Journal of the Indian Mathematical Society	p. 147, §§ 6.5–6
JLMS	Journal of the London Mathematical Society	p. 41, § 2.12
JM	Journal für die reine und angewandte Mathematik	p. 60, § 3.2
MA	Mathematische Annalen	p. 21, § 1.1
MM	Messenger of Mathematics	p. 92, §§ 4.7–8
MTE	Mathematikai és Termész. Értesítő (Budapest)	p. 146, § 6.4
MZ	Mathematische Zeitschrift	p. 63, § 3.7
NA	Nouvelles Annales de Mathématiques	p. 246, § 10.11
OQJ	Quarterly Journal of Mathematics (Oxford)	p. 147, § 6.11
PCPS	Proceedings of the Cambridge Philosophical Society	p. 40, § 2.4
PEMS	Proceedings of the Edinburgh Mathematical Society	p. 245, § 10.3
PLMS	Proceedings of the London Mathematical Society	p. 40, § 2.10(2)
PMF	Prace matematyczno-fizyczne (Warsaw)	p. 60, § 3.2
PNAS	Proceedings of the National Academy of Science (Washington)	p. 281, § 11.10
PTRS	Philosophical Transactions of the Royal Society	p. 40, § 2.5
QJM	Quarterly Journal of Mathematics	p. 63, § 3.8
RP	Rendiconti del Circolo matematico di Palermo	p. 118, § 5.6
TAMS	Transactions of the American Mathematical Society	p. 62, § 3.5(3)
TCPS	Transactions of the Cambridge Philosophical Society	p. 22, § 1.5
TMJ	Tôhoku Mathematical Journal	p. 60, § 3.2
WS	Wiener Sitzungsberichte	p. 246, § 10.11

LIST OF AUTHORS OF ORIGINAL PAPERS AND BOOKS NOT INCLUDED IN THE LIST OF BOOKS

[References to §§ are to the notes at the ends of the chapters, where detailed references will be found]

LIST OF DEFINITIONS

Numbers refer to pages

GENERAL INDEX

Numbers refer to pages

PRINTED IN GREAT BRITAIN AT THE UNIVERSITY PRESS, OXFORD
BY VIVIAN RIDLER, PRINTER TO THE UNIVERSITY